

78-10207



FORMERLY WILLOW RUN LABORATORIES THE UNIVERSITY OF MICHIGAN

CR 151754

Environmental Research Institute of Michigan

P. O. Box 8618

Ann Arbor, Michigan 48107

"Made available under NASA sponsorship
in the interest of early and wide dis-
semination of Earth Resources Survey
Program information and without liability
for any use made thereof."

Progress Report

for

NASA Contract NAS9-15476

ANALYSIS OF SCANNER DATA FOR CROP INVENTORIES

Principal Investigator

Richard F. Nalepka

Task Leaders

Richard J. Kauth

Richard C. Cicone

Peter F. Lambeck

William A. Malila

John E. Colwell

Period Covered

15 March 1978 to 14 June 1978

ERIM Report Number 132400-7-L



(E78-10207) ANALYSIS OF SCANNER DATA FOR
CROP INVENTORIES Progress Report, 15 Mar. -
14 Jun. 1978 (Environmental Research Inst.
of Michigan) 130 p HC A07/MF A01 CSCL 02C

N78-31505

Unclas
G3/43 00207

"Data collected under NASA sponsorship
in the interest of early and wide dis-
semination of Earth Resources Survey
Program information and without liability
for any use made thereof."



FORMERLY WILLOW RUN LABORATORIES THE UNIVERSITY OF MICHIGAN

PREFACE

The following report serves as the technical progress report for Contract NAS9-15476 which is entitled, "Analysis of Scanner Data for Crop Inventories". This report describes the work carried out under a set of six tasks for the period 15 March 1978 through 14 June 1978.

Work on this contract is performed in the Infrared & Optics Division directed by Mr. Richard R. Legault. Dr. Quentin A. Holmes, Head of the Information Systems and Analysis Department is the Project Director for this contract and Mr. Richard F. Nalepka, Head of the Multispectral Analysis Section, is the Principal Investigator.

This contract is part of a comprehensive and continuing program of research at ERIM into earth resources information systems which employ remote sensing of the environment from aircraft and satellites to gather data and which use automatic processing to extract information from the data. The basic objective of this multidisciplinary program is to develop such information systems and practical tools which will provide planners and decision-makers extensive accurate information quickly and economically.

The six tasks reported on for this reporting period include:

- Multisegment Training
- Evaluation of Partitions for Signature Extension
- Wheat Vs. Small Grains
- Forecasting Production of Wheat from Satellite Data
- Prepilot Study of Multicrop Spectral Separability
- Multicrop Labeling Aids

No report is included for the Haze and Soil Correction or the Evaluation of Multitemporal Classifiers tasks since, for this quarter, the sponsor has redirected our efforts to multicrop associated activities.



TASK 1
MULTISEGMENT TRAINING
(R. Kauth and W. Richardson)

1.1 OBJECTIVE

The objective of Task 1 is to develop a sampling strategy for selecting training data, applicable to proportion estimation over a wide region. The main requisites of that strategy are that it produces a representative sample and that the training sample size is small compared to the total area to be classified.

1.2 APPROACH

1. Create a conceptual basis for the problem of training in a large scale remote sensing system, incorporating the inputs from UCB, LARS, and other ERIM tasks, and consistent with LACIE operational constraints.
2. Within this framework, propose a detailed methodology for training selection.
3. Demonstrate the selection methodology in an intermediate scale exercise over a partition containing from 15 to 30 sample segments from which 5 to 10 segments are selected for training and for which a wheat proportion estimate is made.
4. Incorporate both multitemporal and across partition signature extension capability into the final procedure.
5. Incorporate the capability to work with incomplete sets of multitemporal data and to optimize selection to make estimates at several times during the growing season.

1.3 PROGRESS

During the last quarter a baseline version of Procedure B was provided to Task 2 of this contract for testing. The approach to the missing acquisition capability was further clarified and coding modifications needed were identified. Coding of the missing acquisition capability is about 90% complete. The problem of defining a composite procedure combining desirable aspects of both Procedure B and JSC's Procedure 1 was considered. No final design was defined but some suggestions were included in the SR&T quarterly review, June 12 - 16. ERIM personnel took part in the Procedure 2 design review, June 16, 1978.

The major effort during the quarter was in the development and exercising of diagnostic tools and procedures to measure the performance of components of Procedure B. The tools developed are similar to some already developed at JSC to measure component performance of Procedure 1. Note that the tests planned on the baseline version of Procedure B under Task 2 are tests of global performance compared to other approaches. The component performance tests being developed and carried out under this task are for the purpose of identifying and isolating the sources of variance and bias in Procedure B and of establishing optimal parameter settings.

1.4 TECHNICAL DISCUSSION

The technical discussion will be limited to the development and exercise of the component performance tests. Some introductory discussion is followed by the material presented at the quarterly review.

1.4.1 COMPONENTS AND PERFORMANCE MEASURES

Currently, three major components of Procedure B are being examined:

1. Spectral/Ancillary Data Stratification (B-CLUSTERING)
2. Training Segment Selection
3. Training Blob Selection

The performance measures which are being used to evaluate these components are the variance, the bias, and the number of training segments required to represent the entire set of segments.

1.4.2 SPECTRAL/ANCILLARY DATA STRATIFICATION

The main performance measures to be used in evaluating the performance of B-CLUSTERING are the reduction in variance of the proportion estimate due to the clustering and the number of training segments required to achieve 90% of maximum "value" in segment selection.

The formula for reduction of variance due to stratification is

$$f = \frac{\sum N_i p_i (1 - p_i)}{N p (1 - p)}$$

where the sum is over all B-clusters

N_i is the number of pixels in the i^{th} B-cluster and N is the total number,

p_i is the proportion of wheat in the i^{th} B-cluster and p is the average proportion of wheat,

f is a number which ranges from 0 to 1, and is better if it is smaller.

The reduction of variance is computed using ground truth sites so that the quantities N_i and p_i are known.

The reduction in variance of the B-CLUSTERING is affected by the spectral channels used, the ancillary variables used, and in particular by the relative weights used. As the weight on the ancillary variables is increased, the stratification becomes better but the number of segments required to represent the entire partition increases. Hence the best choice of ancillary and spectral variables is the one with the

smallest f value which achieves a given percent of representation with a certain fraction of the segments used for training.

1.4.3 TRAINING SEGMENT SELECTION

The performance measures to be used in evaluating the training segment selection procedure are the variance and bias of the partition-wide proportion estimate and the individual segment proportion estimates. The parameters which are varied are the number of segments chosen for training and the number of blobs chosen for training. In particular, it is of interest to know whether the process of selecting segments intrinsically introduces a bias into the procedure. If so, it may be necessary to randomly select samples from all segments. Even if this turns out to be the case, Procedure B may have sampling advantages over a single segment procedure.

1.4.4 TRAINING BLOB SELECTION

The performance measures for training blob selection are again the variance and bias of partition-wide proportion estimates and of individual segment estimates. The parameters which can be varied are the minimum number of internal pixels and the shape factor (compactness) in blobs chosen for training.

In order to measure variance and bias, it is necessary to create random replicates of the selection process. This has been done by making different random choices of the blobs actually used in the procedure.

1.4.5 QUARTERLY REVIEW PRESENTATION

The following presentation includes recommendations and future plans.

OUTLINE

- INTRODUCTION
- STATUS
- EXPERIMENTAL RESULTS
- CONCLUSIONS
- RECOMMENDATIONS

TASK 1

INTRODUCTION TO PROCEDURE B

- PROCEDURE B IS A TECHNIQUE OF PROPORTION ESTIMATION WHICH TELLS AN ANALYST WHICH SCENE ELEMENTS TO LABEL, AND THEN USES THOSE LABELS IN AN UNBIASED WAY TO PRODUCE A PROPORTION ESTIMATE FOR THE SCENE
- FUNDAMENTAL CONCEPT SIMILAR TO PROCEDURE 1
 - STRATIFIED SAMPLING USED FOR "BIAS CORRECTION"
 - SAMPLING APPROXIMATELY PROPORTIONAL TO SIZE OF STRATA MINIMIZES VARIANCE
 - STRATA FORMED BY "CLASSIFICATION" OF SAMPLES USING A CLUSTERING ALGORITHM OPERATING ON A FEATURE VECTOR
- MAJOR DIFFERENCES IN CONCEPT
 - PREPROCESSING/FEATURE EXTRACTION
 - NUMBER OF STRATA
 - NUMBER OF SEGMENTS

TASK 1

PREPROCESSING/FEATURE EXTRACTION

- DATA SCREENING
- SUN ANGLE AND HAZE CORRECTION
- FEATURES USED
 - SPECTRAL: BRIGHTNESS/GREENNESS
 - TEMPORAL: 2 OR 3 BIOWINDOWS SIMULTANEOUSLY
 - SPATIAL: CONTIGUOUS GROUPS OF SPECTRALLY-TEMPORALLY SIMILAR PIXELS (BLOBS)
 - ANCILLARY DATA:
PARAMETERS SUCH AS VIEW ANGLE, CROP CALENDAR,
AVAILABLE SOIL MOISTURE, LATITUDE AND LONGITUDE,
WHICH VARY FROM SEGMENT TO SEGMENT

TASK 1



STRATIFICATION PROCEDURE

- CLUSTER TOGETHER DEFINED SPECTRAL-SPATIAL-TEMPORAL FEATURES
 - MULTISEGMENT ($\sim 1/2$ STATE OF KANSAS)
 - ANCILLARY DATA (VALUES SPECIFIC TO EACH SEGMENT)
 - MULTIPLE STRATA (B-CLUSTERS)

TRAINING SELECTION PROCEDURE

- PROCEDURE HAS EVOLVED TO AN ALLOCATION PROCEDURE FOR STRATIFIED SAMPLING
- ATTEMPT TO SELECT A SUBSET OF SEGMENTS WHICH WILL PROVIDE A SUFFICIENT SUPPLY OF SAMPLES FOR LABELING WITHIN EACH STRATA

TASK 1

PROPORTION ESTIMATION PROCEDURE

- ESTIMATED PROPORTION IS THE WEIGHTED AVERAGE OF THE ESTIMATED PROPORTIONS IN ALL THE B-CLUSTERS, WHERE THE WEIGHTS ARE THE NUMBER OF PIXELS IN EACH B-CLUSTER
- THE ESTIMATED PROPORTION IN EACH B-CLUSTER IS THE WEIGHTED AVERAGE OF THE ESTIMATED PROPORTIONS OF EACH BLOB USED FOR LABELING, WHERE THE WEIGHTS ARE THE NUMBERS OF PIXELS IN EACH BLOB
- THE ESTIMATED PROPORTIONS IN EACH BLOB COME FROM ANALYSTS (GROUND TRUTH) IN PRODUCTION (RESEARCH) VERSION

1-10

TASK 1



PROCEDURE B PROPORTION ESTIMATION FORMULA

$$\hat{p} = \frac{\sum_{I=1}^M N_I \hat{p}_I}{\sum_{I=1}^M N_I}$$

WHERE

N_I IS THE TOTAL NUMBER OF PIXELS IN THE I^{TH} BCLUSTER

$$\hat{p}_I = \frac{\sum_{K \in T(I)} B_{IK} \hat{p}_{IK}}{\sum_{K \in T(I)} B_{IK}}$$

AND WHERE

$T(I)$ IS THE SET OF BLOB CHOSEN FOR TRAINING THE I^{TH} BCLUSTER

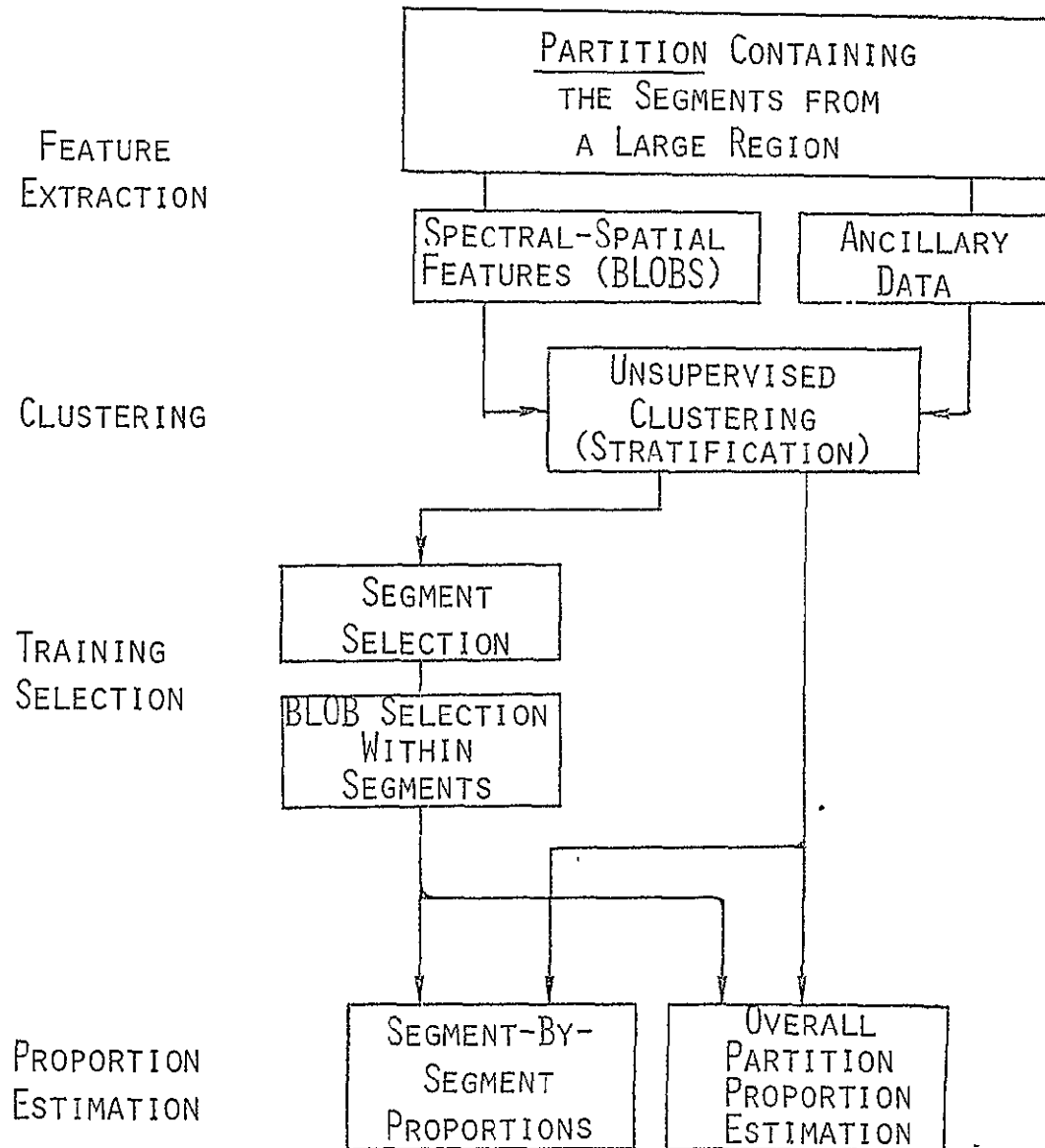
B_{IK} IS THE NUMBER OF PIXELS IN THE IK^{TH} BLOB USED FOR TRAINING

$\hat{p}_{IK}$ IS THE ESTIMATED PROPORTION FOR THE IK^{TH} BLOB

TASK 1



PROCEDURE B DATA FLOW



1-12

TASK 1

STATUS OF PROCEDURE B DEVELOPMENT

- HAVE ESTABLISHED A BASELINE PROCEDURE FOR T&E
- HAVE INITIATED EXPERIMENTS TO MEASURE PERFORMANCE OF THE COMPONENTS OF PROCEDURE B

TASK 1



EXPERIMENTS

- OBJECTIVE
 - DETERMINE OPTIMAL PARAMETER SETTINGS FOR COMPONENTS OF PROCEDURE B
- APPROACH
 - MEASURE COMPONENT PERFORMANCE AS FUNCTION OF PARAMETER VALUES
- MEASURES OF COMPONENT PERFORMANCE
 - VARIANCE REDUCTION FACTOR (R.V.)
 - SAMPLING BIAS
 - SAMPLING VARIANCE



TASK 1

COMPONENTS BEING EXAMINED

- SPECTRAL/SPATIAL STRATIFICATION (BLOB)
- SPECTRAL/ANCILLARY DATA STRATIFICATION (B-CLUSTER)
- TRAINING SEGMENT SELECTION
- TRAINING BLOB SELECTION
- PROPORTION ESTIMATION ALGORITHM

TASK 1

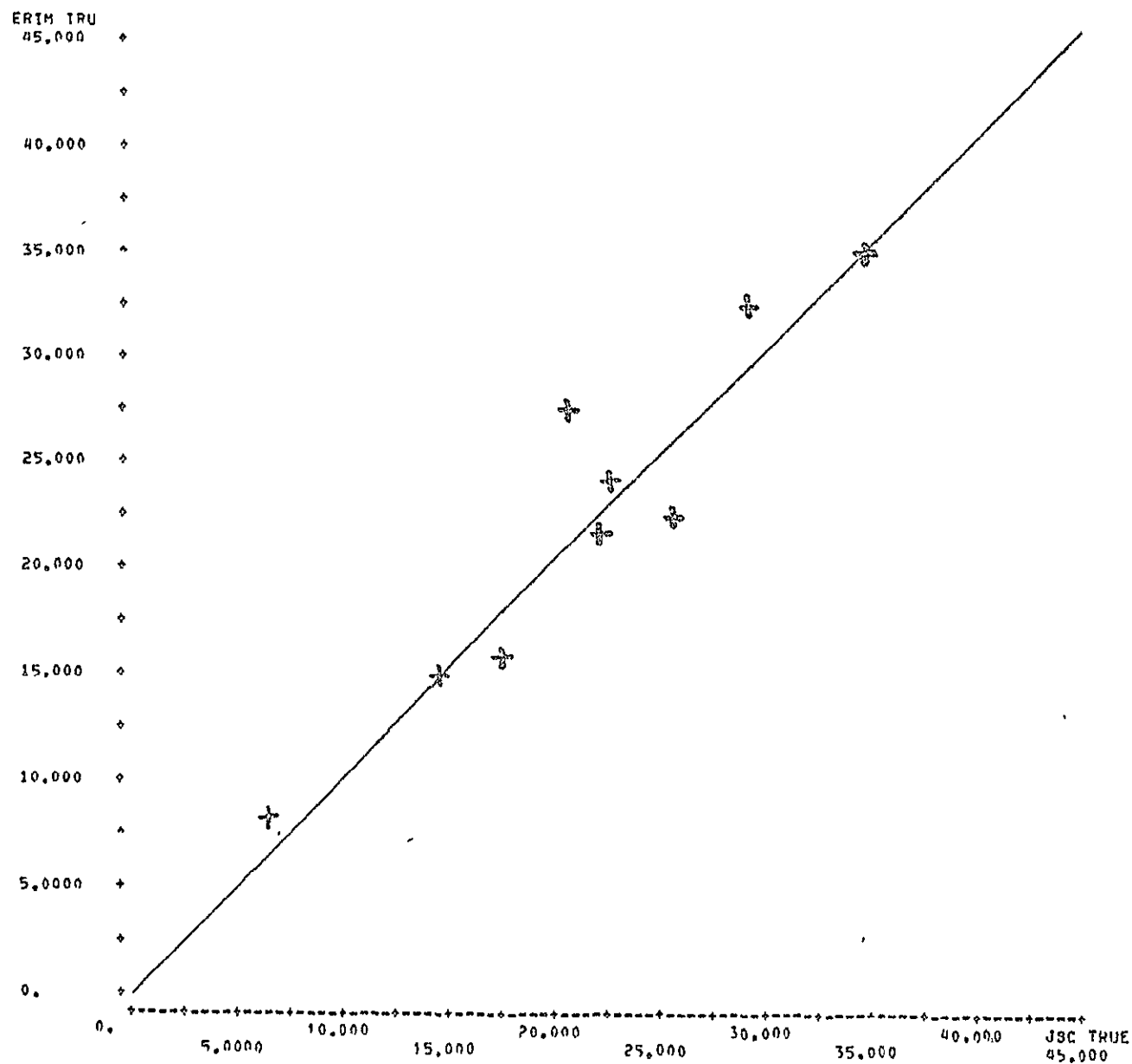
SELECTION/PROPORTION ESTIMATION

- PARAMETERS TO VARY OR EFFECTS TO EVALUATE
 - SEGMENT SELECTION
 - : NUMBER OF SEGMENTS CHOSEN
 - : VARIATION DUE TO CHOICE
 - BLOB SELECTION
 - : NUMBER OF BLOBS CHOSEN
 - : VARIATION DUE TO CHOICE
 - PROPORTION ESTIMATION
 - : ALGORITHM DETAIL
- PERFORMANCE MEASURES ARE BIAS AND VARIANCE OF PROPORTION ESTIMATE FOR THE ENTIRE GROUP OF SEGMENTS

TASK 1



SCATTER PLOT OF ERIM MANUAL GROUND TRUTH AND JSC GROUND TRUTH FOR 9 SEGMENTS



TASK 1



22.4
21.3

ERIM TRUE MEAN

JSC TRUE MEAN

1-18

ORIGINAL PAGE IS
OF POOR QUALITY

TASK 1

Σ ERIM

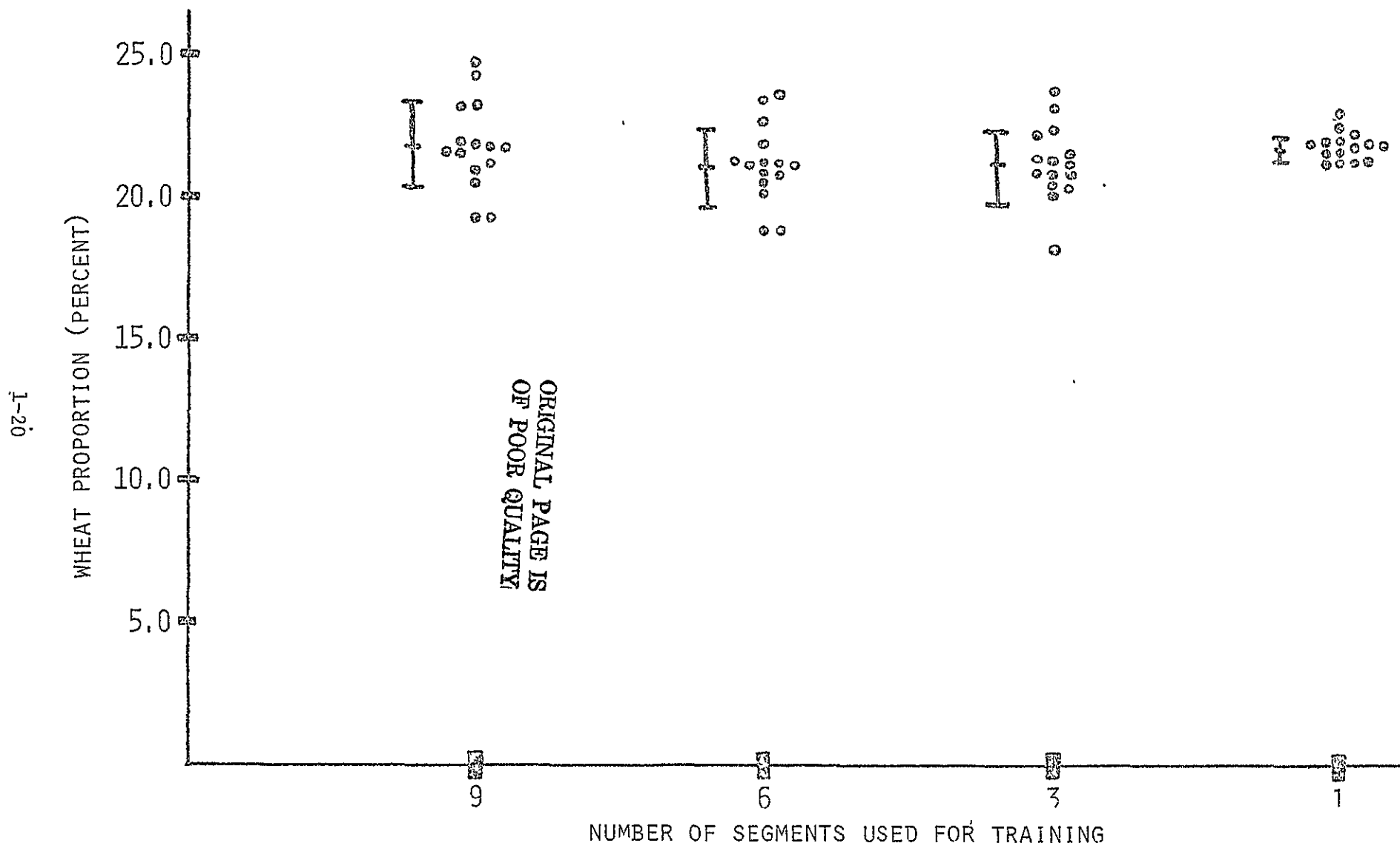
21,9

9 SEGMENT
3000 BLOB
REFERENCE

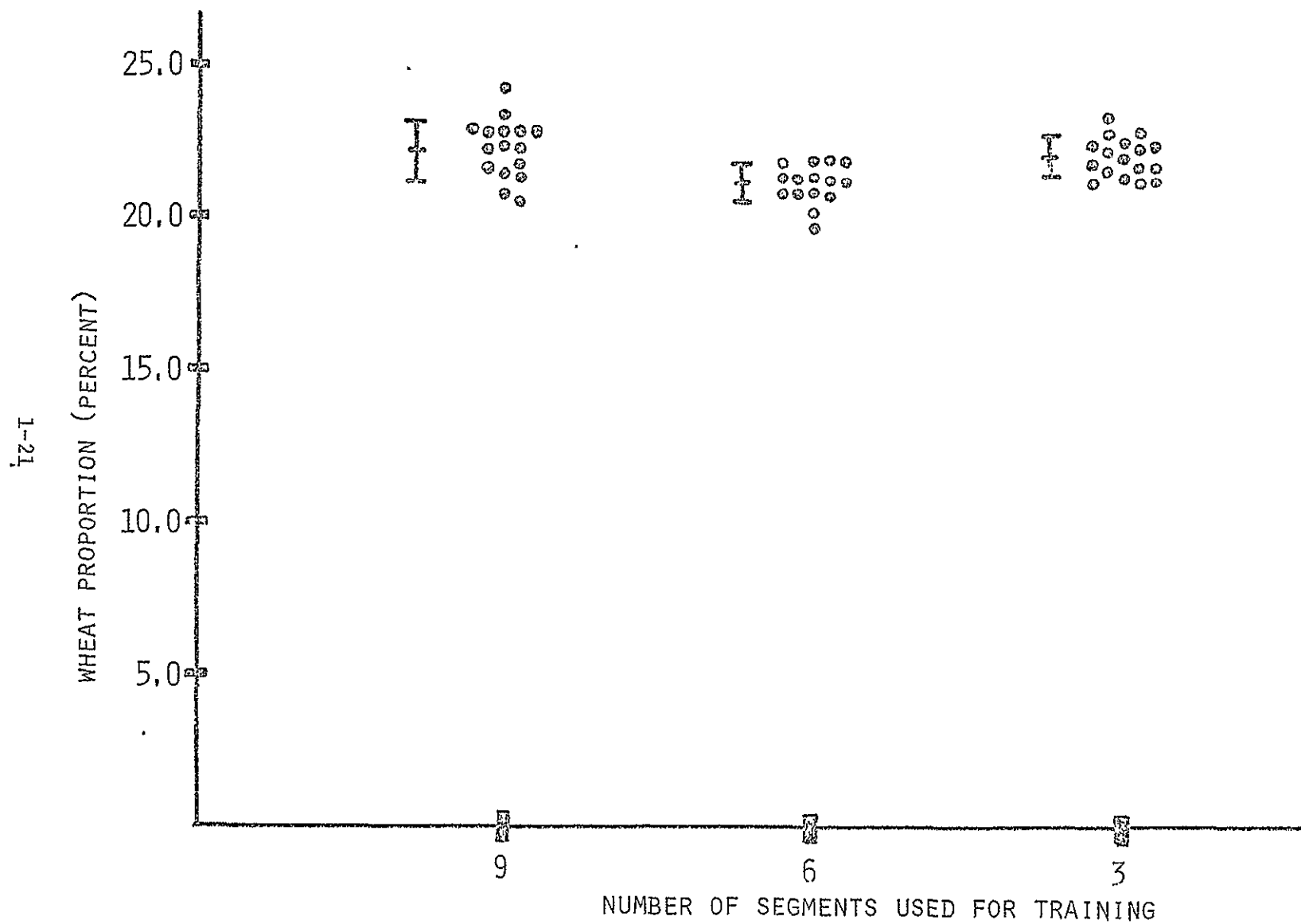
61-1-19

TASK 1

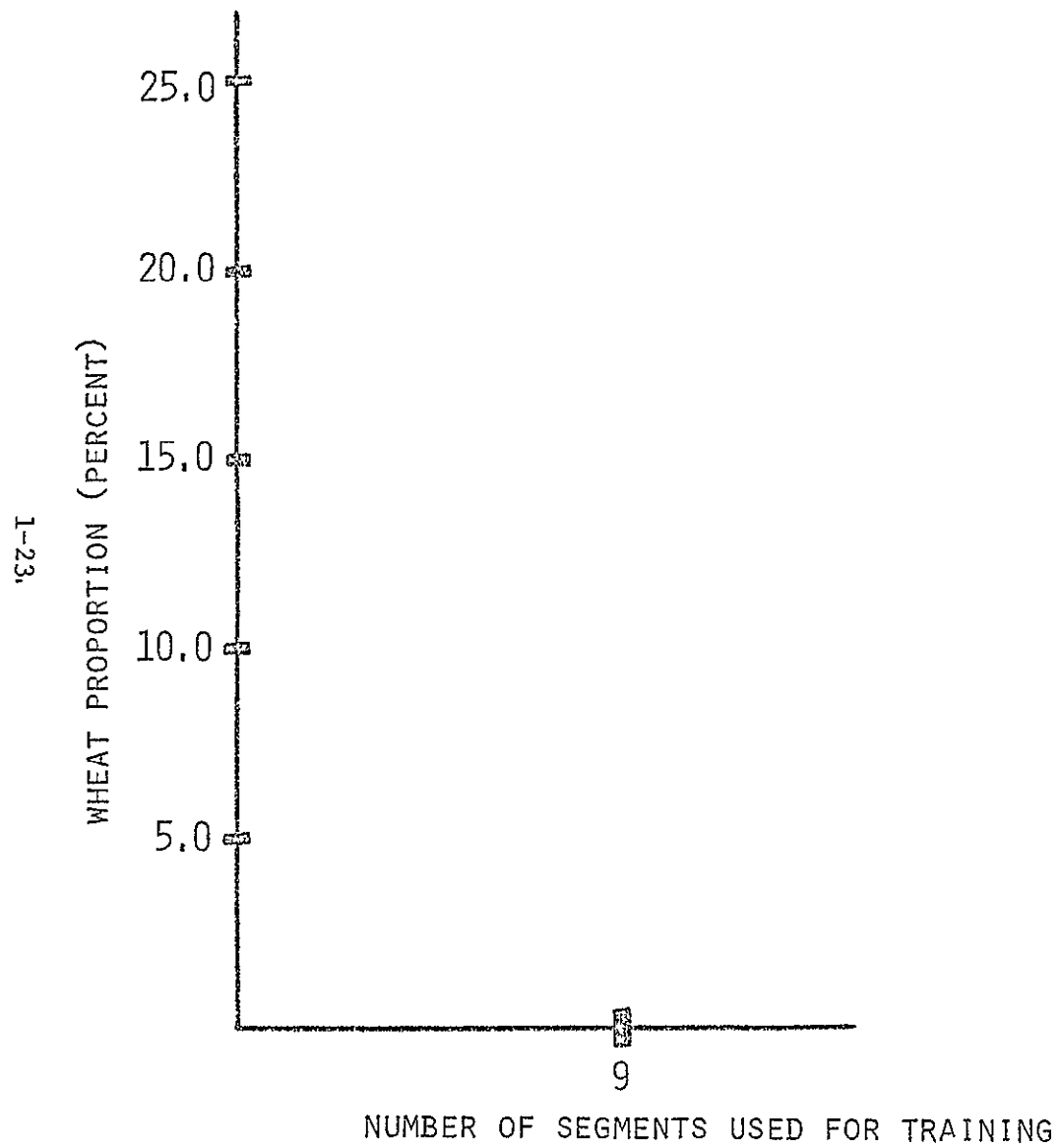




PROCEDURE B: PROPORTION ESTIMATES FOR 9 SEGMENTS WITH APPROXIMATELY 300 BLOBS
SELECTED FOR TRAINING AND WITH NO ANCILLARY VARIABLE.

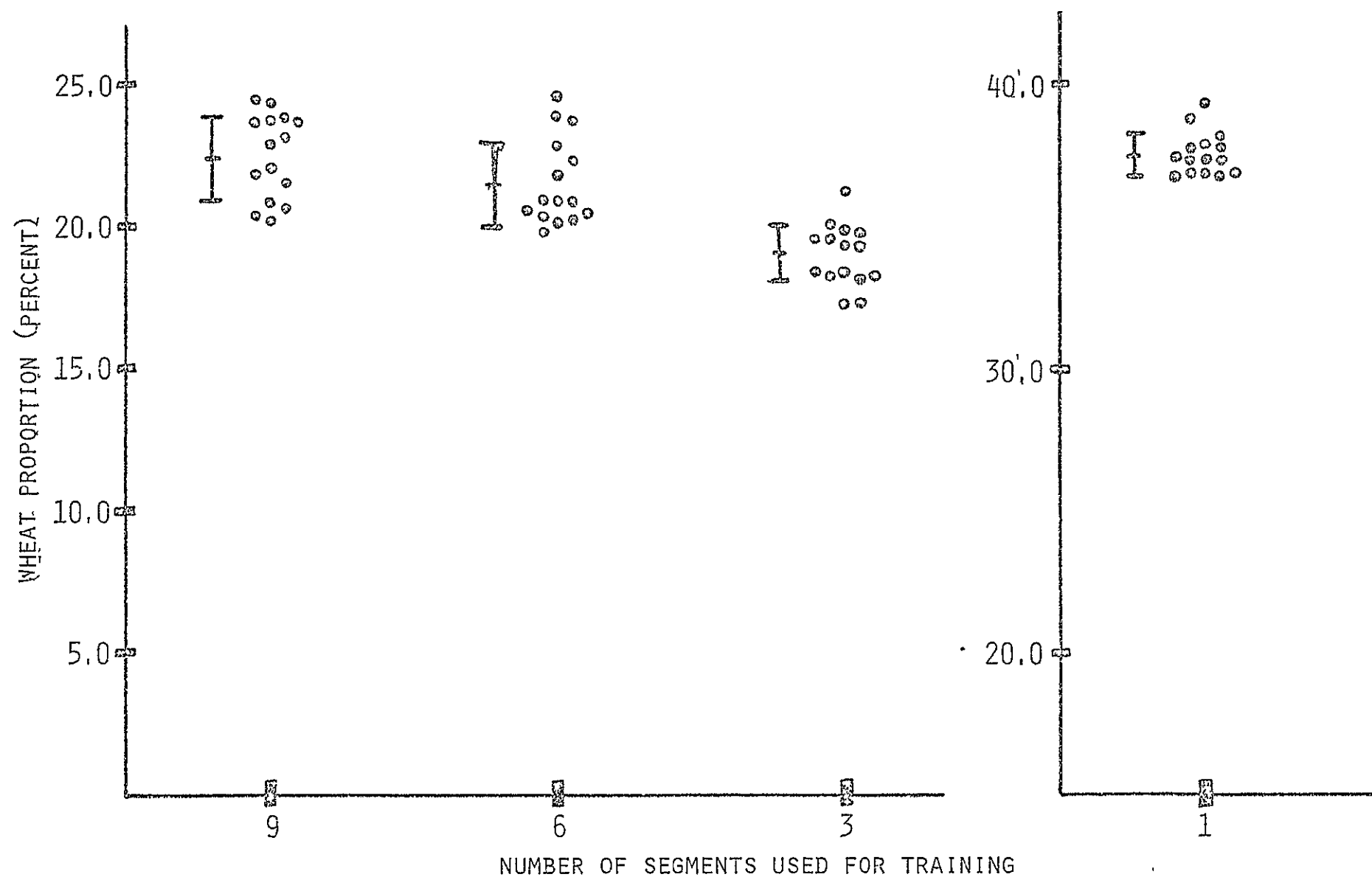


PROCEDURE B: PROPORTION ESTIMATES FOR 9 SEGMENTS WITH APPROXIMATELY 600 BLOBS
SELECTED FOR TRAINING AND WITH NO ANCILLARY VARIABLE.

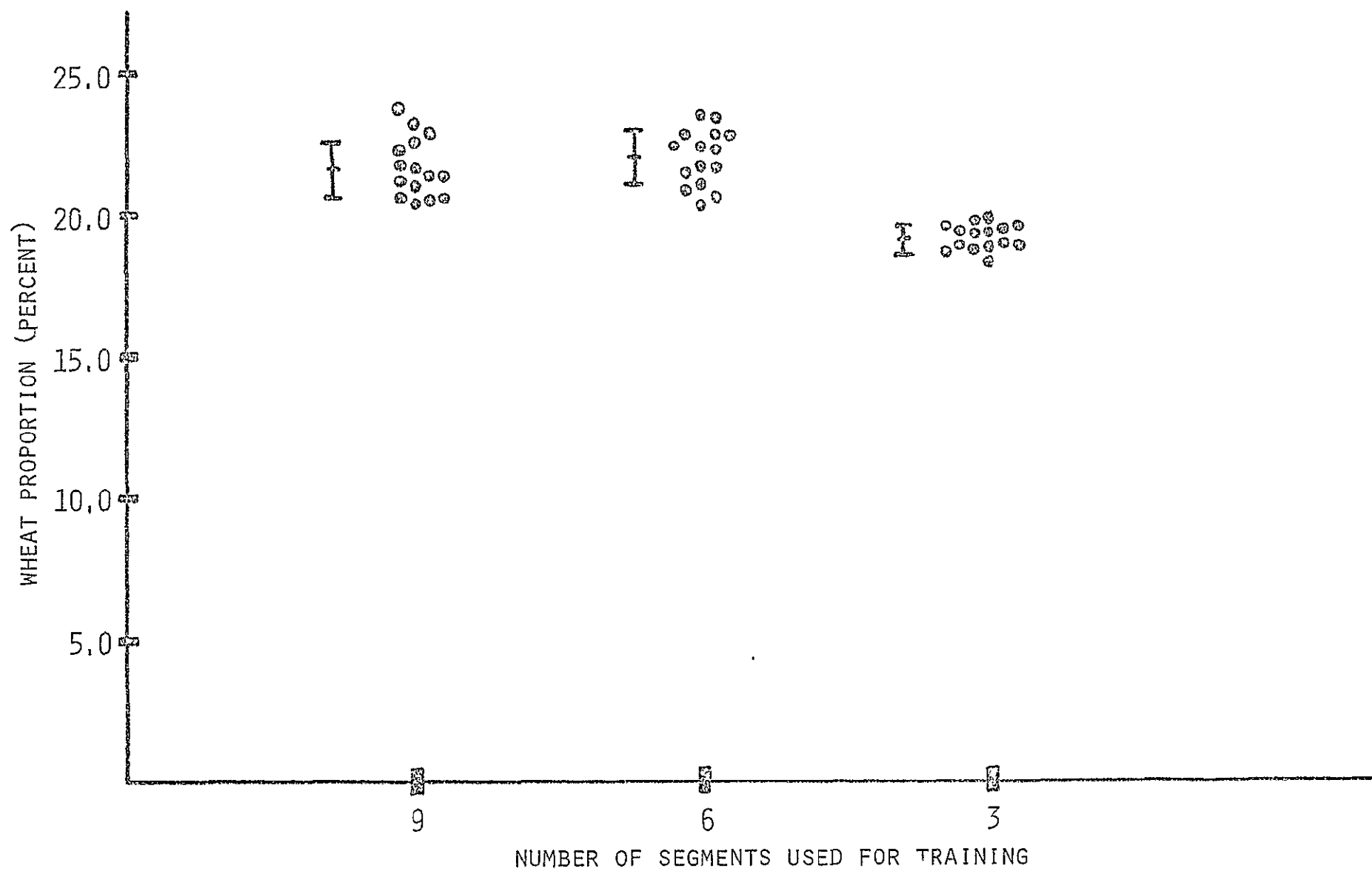


PROCEDURE B: PROPORTION ESTIMATES FOR 9 SEGMENTS WITH APPROXIMATELY 3000 BLOBS
SELECTED FOR TRAINING AND WITH NO ANCILLARY VARIABLE.

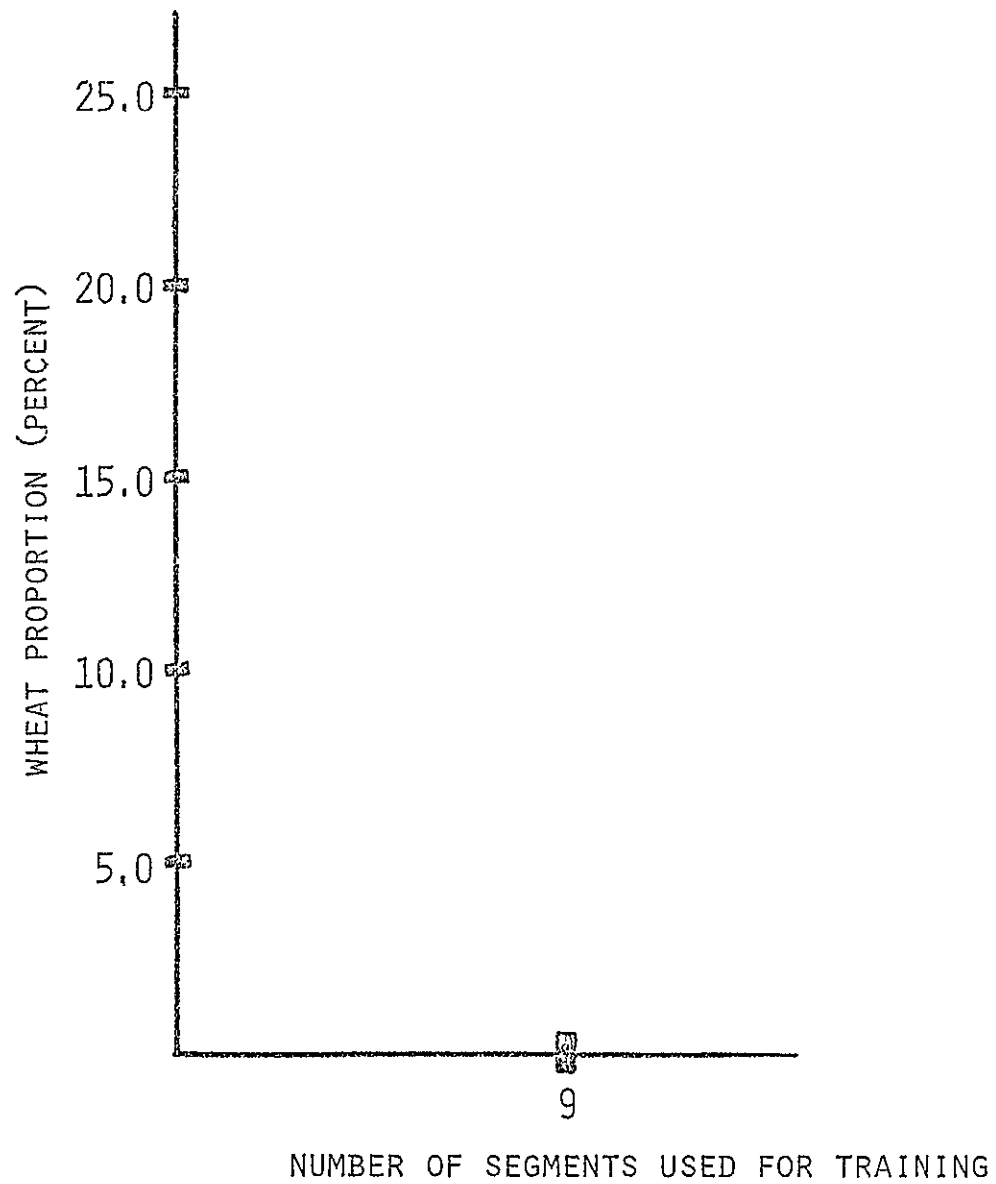
1-24.



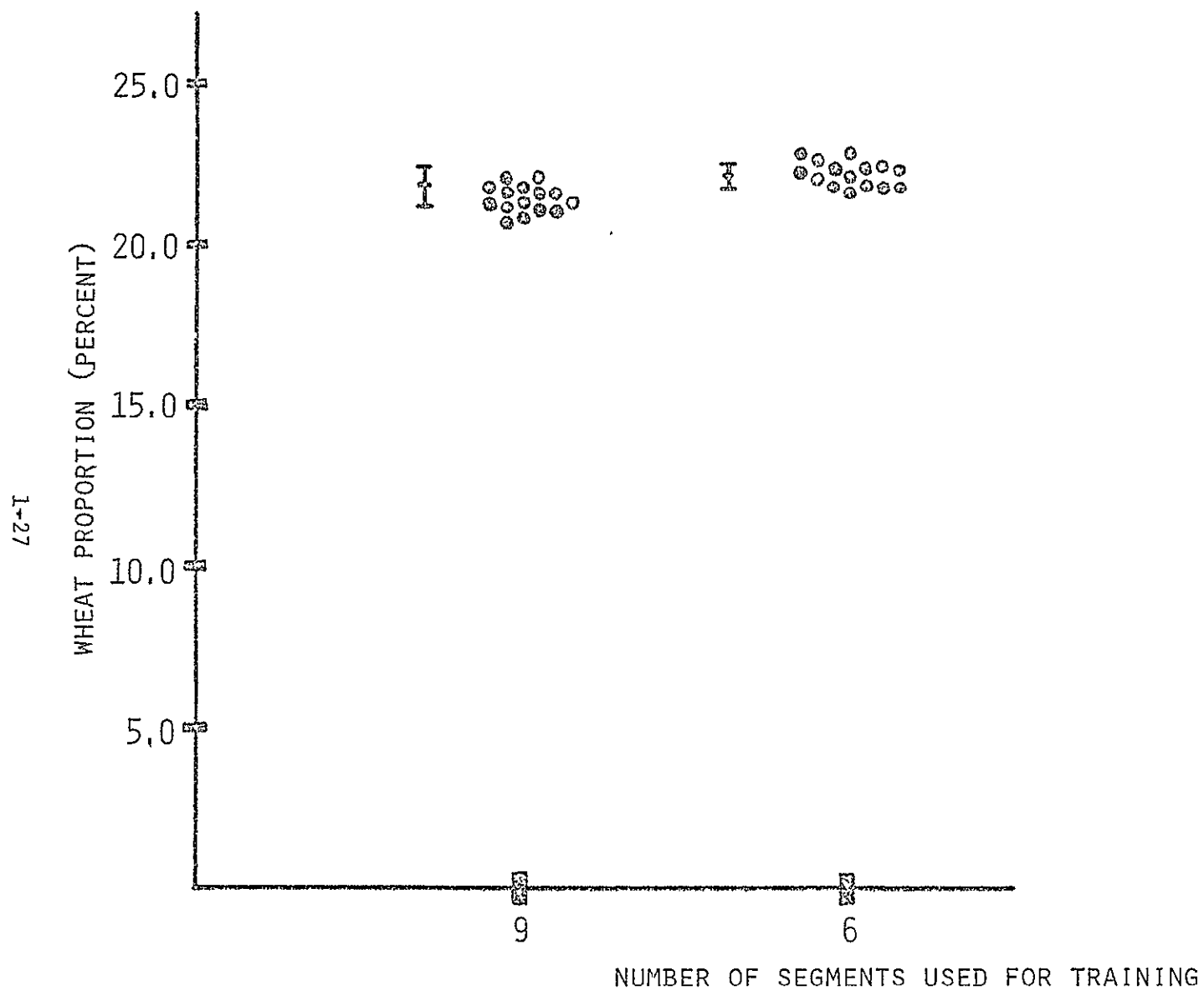
PROCEDURE B: PROPORTION ESTIMATES FOR 9 SEGMENTS WITH APPROXIMATELY 300 BLOBS
SELECTED FOR TRAINING AND WITH NOVEMBER SOIL MOISTURE.



PROCEDURE B: PROPORTION ESTIMATES FOR 9 SEGMENTS WITH APPROXIMATELY 600 BLOBS
SELECTED FOR TRAINING AND WITH NOVEMBER SOIL MOISTURE.



PROCEDURE B: PROPORTION ESTIMATES FOR 9 SEGMENTS WITH APPROXIMATELY 3000 BLOBS
SELECTED FOR TRAINING WITH NOVEMBER SOIL MOISTURE.



PROCEDURE B: PROPORTION ESTIMATES FOR 9 SEGMENTS WITH APPROXIMATELY 1500 BLOBS
SELECTED FOR TRAINING AND WITH NOVEMBER SOIL MOISTURE.

SPECTRAL/ANCILLARY DATA STRATIFICATION

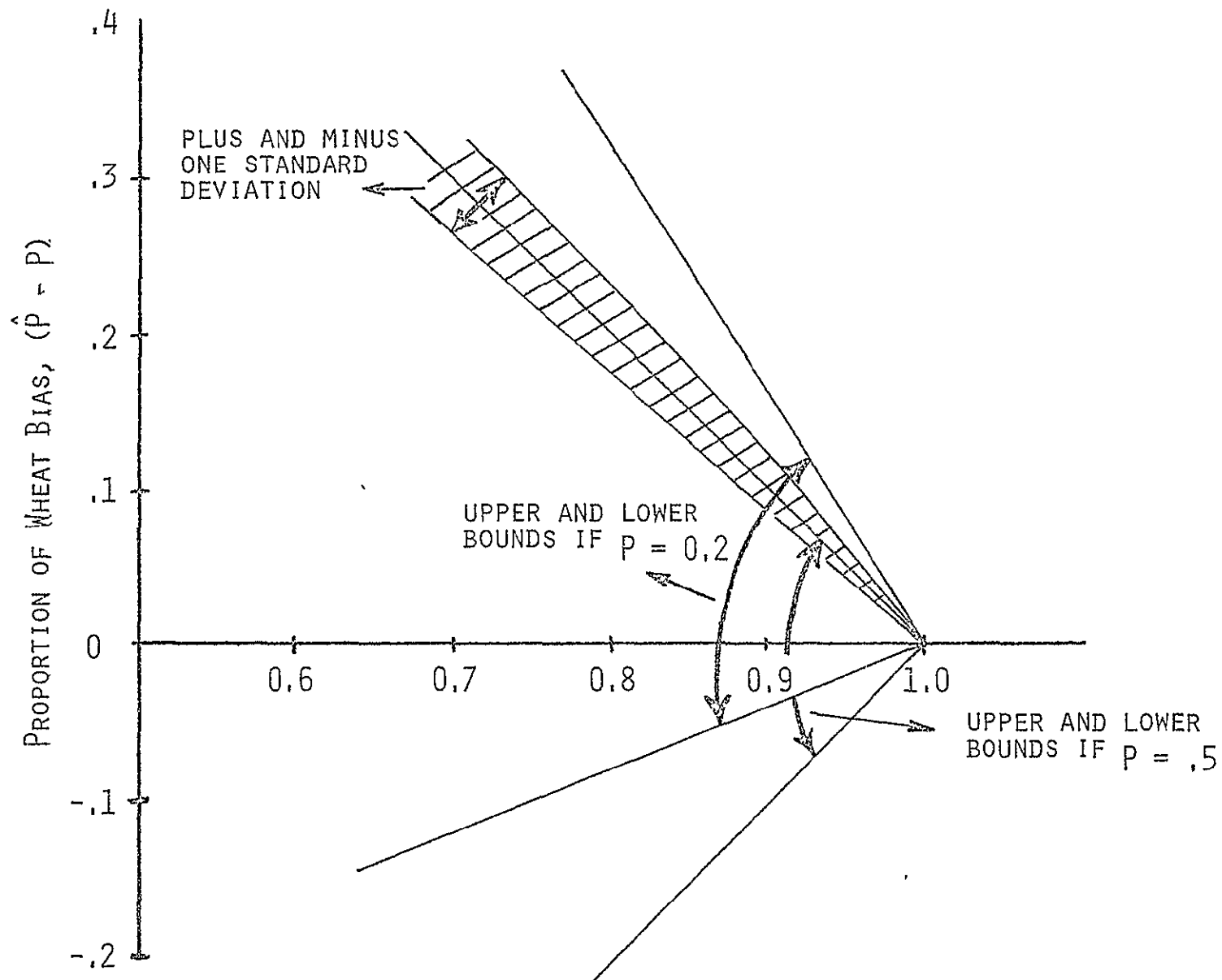
- PARAMETERS TO BE VARIED
 - SPECTRAL VARIABLES USED
 - ANCILLARY VARIABLES USED
 - RELATIVE WEIGHTING
- PERFORMANCE MEASURE IS REDUCTION OF VARIANCE DUE TO STRATIFICATION

$$F = \frac{\sum N_i P_i (1-P_i)}{N P(1-P)}$$

TASK 1

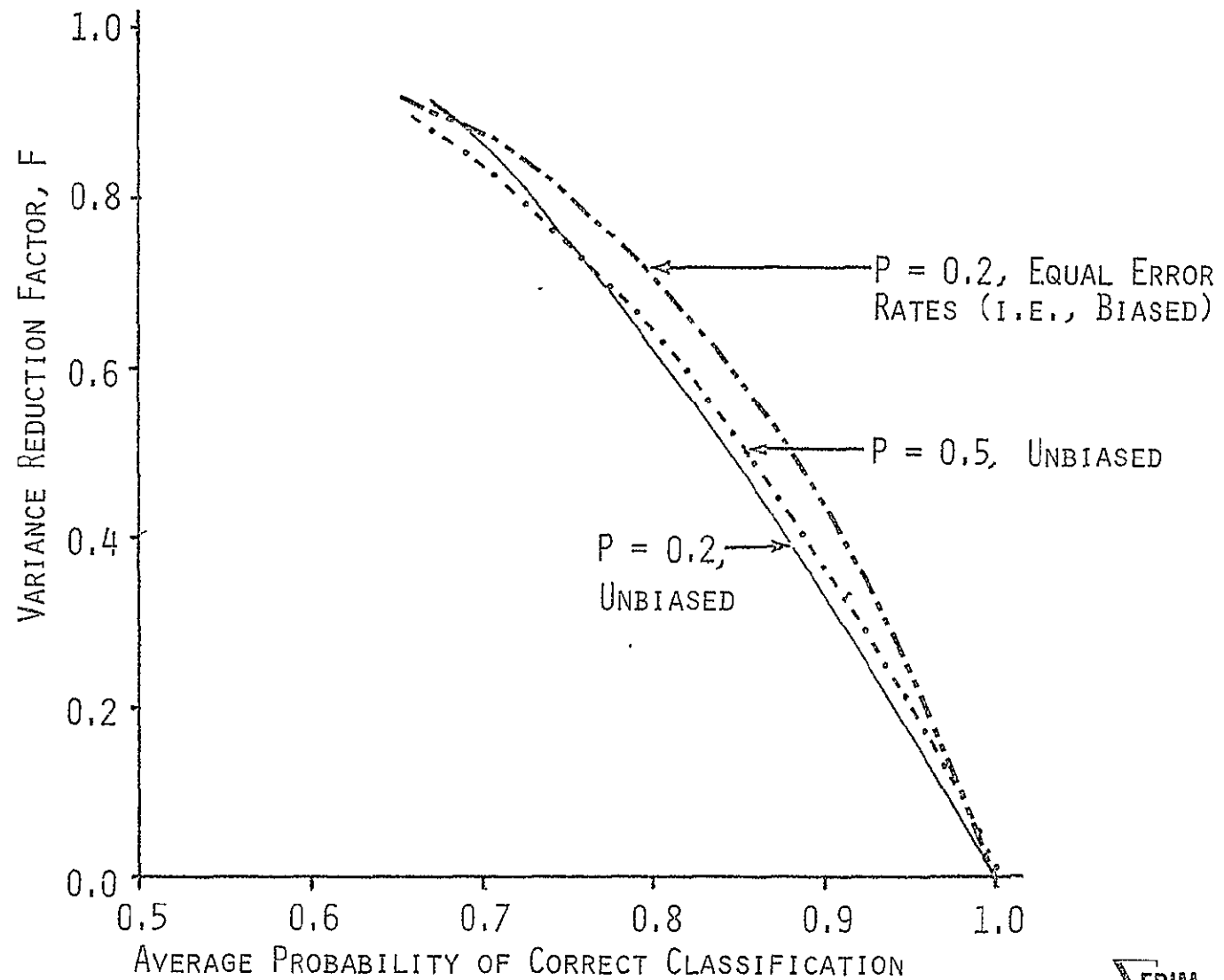


BOUNDS ON CLASSIFIER PROPORTION ESTIMATION BIAS AS A FUNCTION OF AVERAGE PROBABILITY OF CORRECT CLASSIFICATION



1-29

VARIANCE REDUCTION FACTOR FOR SEVERAL CLASSIFIERS WHEN THE TRUE PROPORTION OF WHEAT IS P



RELATIVE WEIGHTING ON 6 SPECTRAL VARIABLES (GREENESS
AND BRIGHTNESS FOR 3 BIOWINDOWS)

- VARY ALL 6 VALUES
- NORMALIZE TO 90 BCLUSTERS
- METHOD OF STEEPEST DESCENT
- 2 WAY R.V.
- FOUND BEST OPERATING POINT WHEN VARIABLES
WERE WEIGHTED INVERSELY TO THEIR RANGE
- 1 WAY R.V. = ,539
- 2 WAY R.V. = ,406

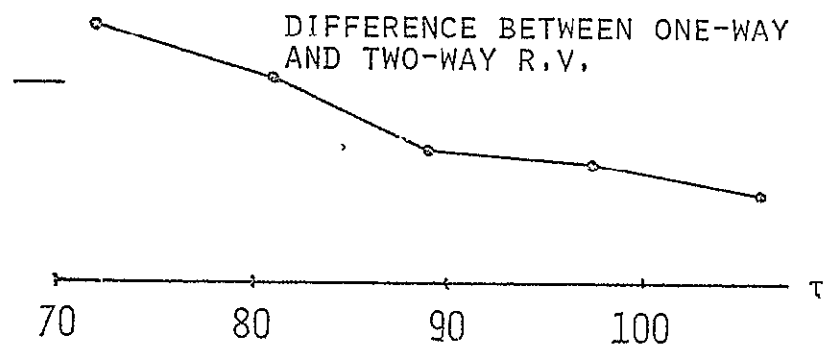
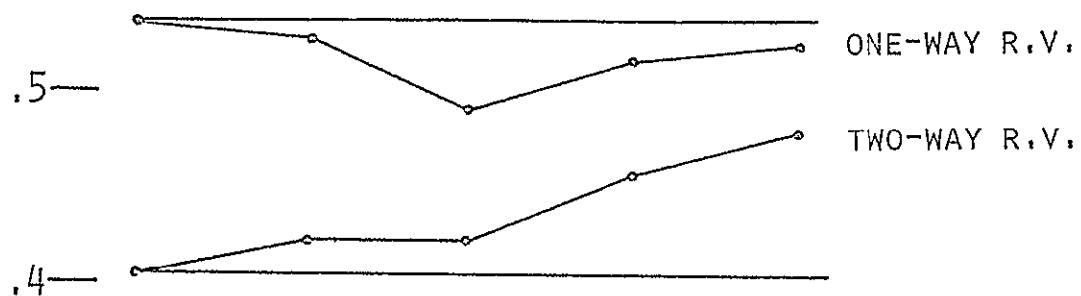
TASK 1

3 GREEN VARIABLES COMPARED WITH 3 GREEN AND 3 BRIGHTNESS VARIABLES

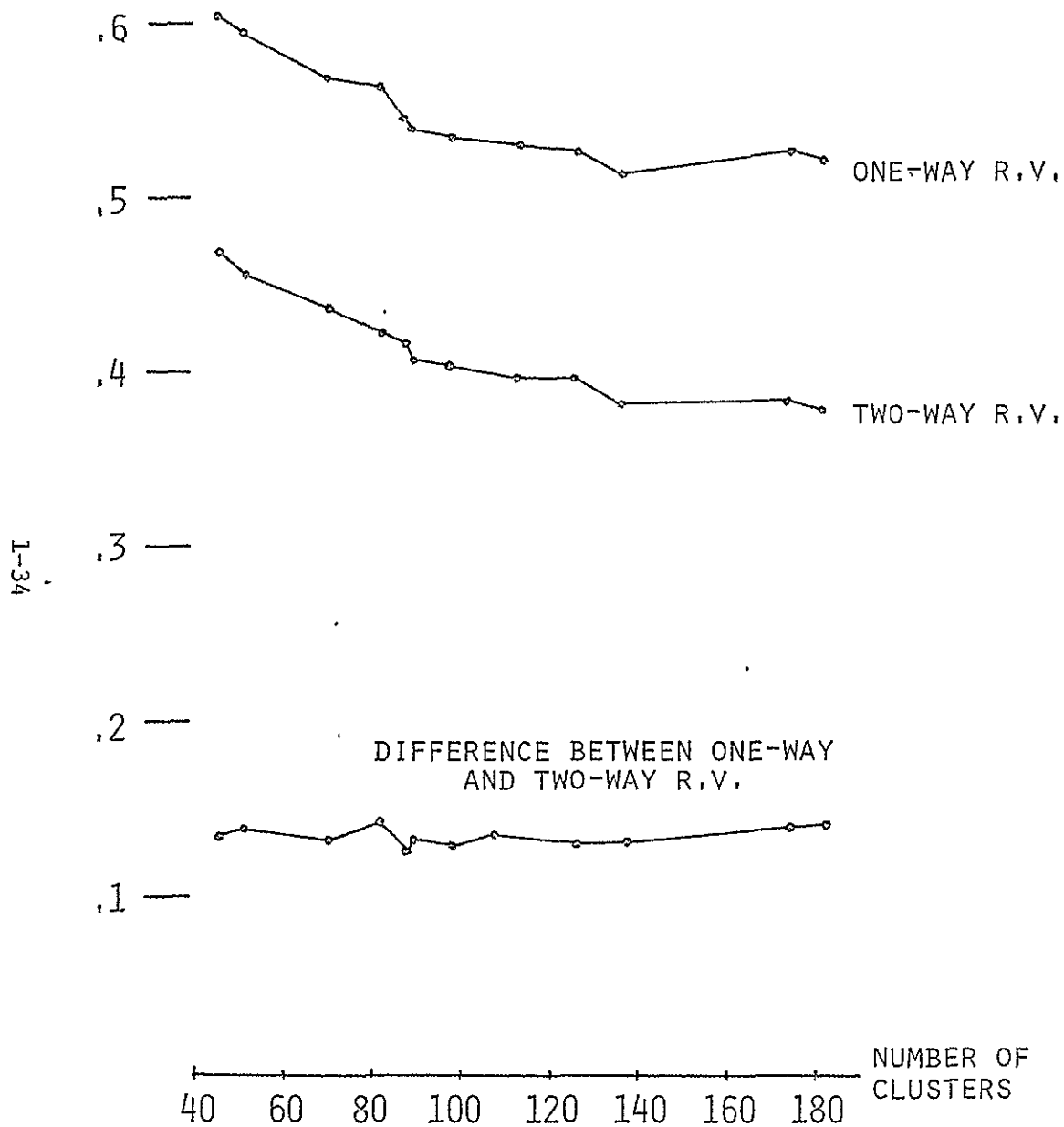
	3G + 3B	3G
1 WAY R.V.	.539	.622
2 WAY R.V.	.406	.430

TASK 1

The logo consists of a large Greek letter sigma (Σ) followed by the word "ERIM" in a bold, sans-serif font.



REDUCTION OF VARIANCE FACTOR DUE TO CLUSTERING
 NOVEMBER 1 SOIL MOISTURE
 ALONG WITH THE SPECTRAL VARIABLES,
 τ MEASURES THE WEIGHT PUT ON THE ANCILLARY
 VARIABLE. $\tau = 72$ CORRESPONDS TO ZERO WEIGHT.

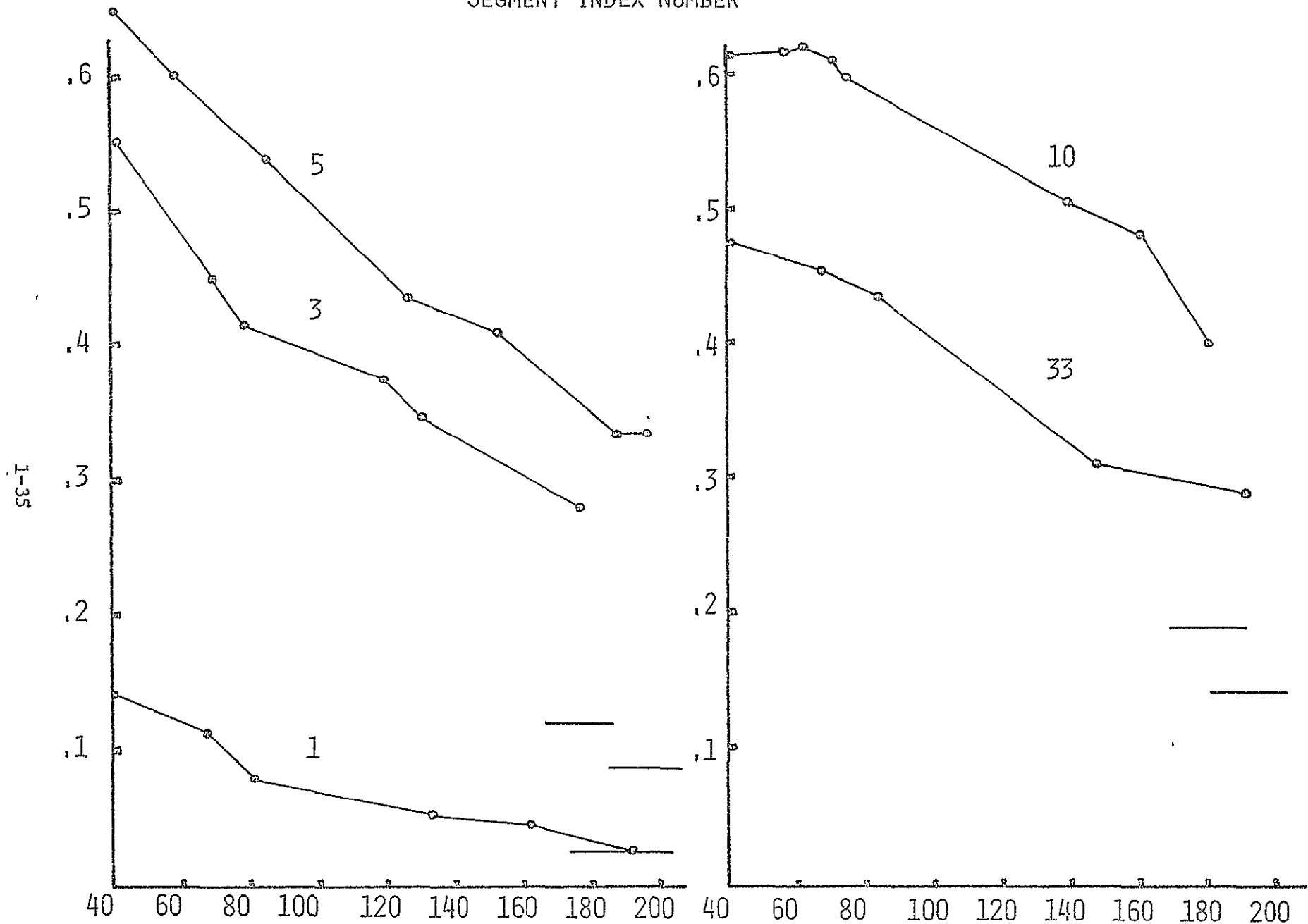


REDUCTION OF VARIANCE FACTOR DUE TO
CLUSTERING
6 SPECTRAL VARIABLES
AS A FUNCTION OF THE NUMBER OF CLUSTERS

TASK 1



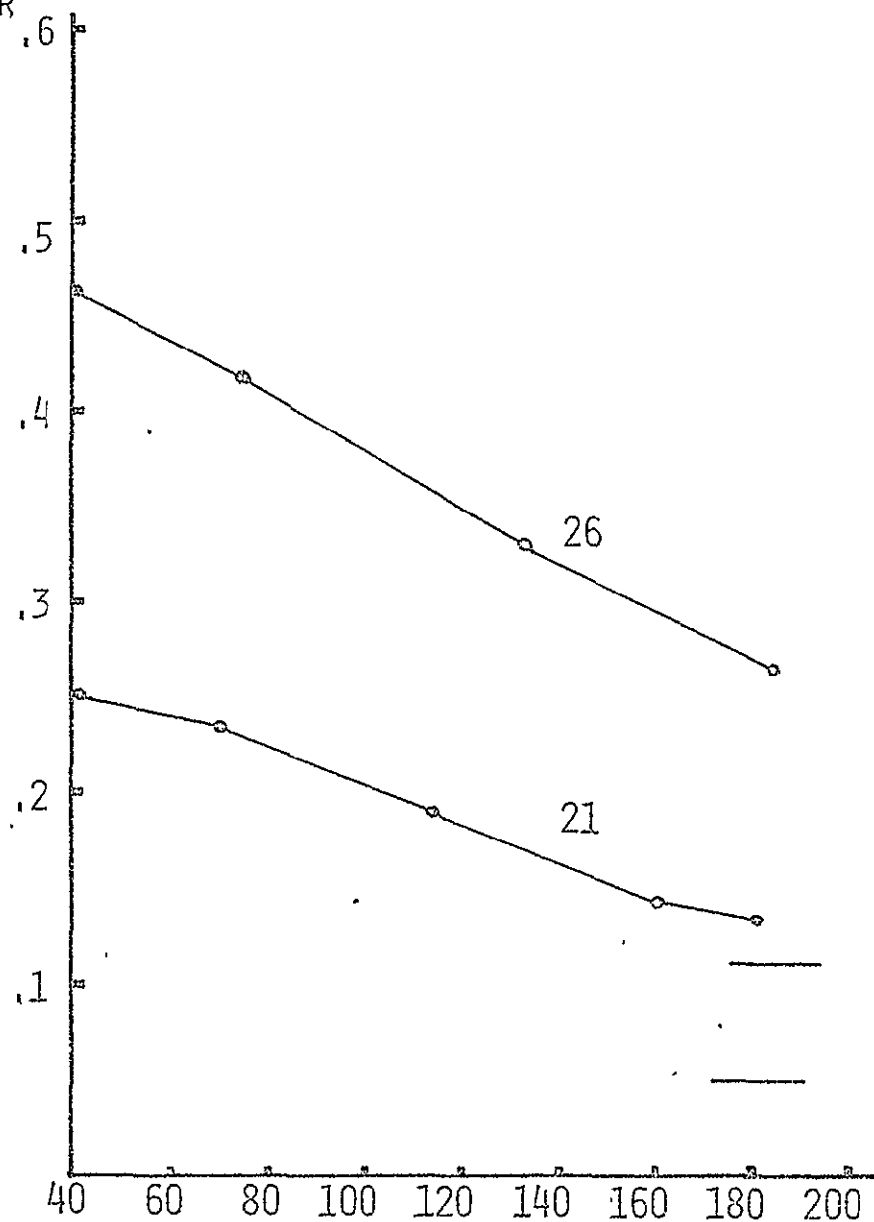
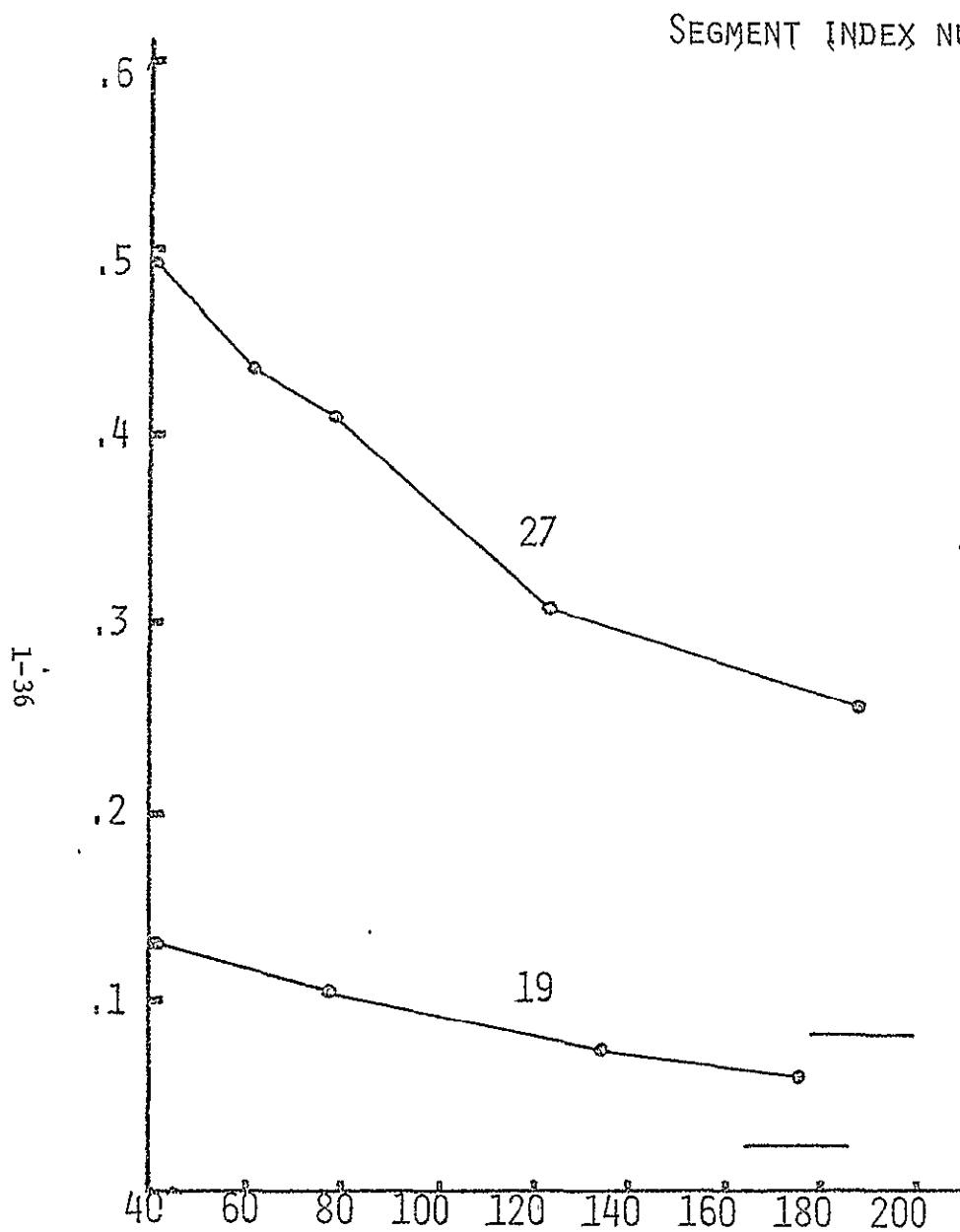
SEGMENT INDEX NUMBER



VARIANCE REDUCTION FACTOR AS A FUNCTION OF NUMBER OF RCLUSTERS FOR SINGLE SEGMENTS. VARIANCE REDUCTION FACTOR DUE TO BLOBBING IS SHOWN AS A LOWER BOUND, SLIDE 2 OF 2,

TASK 1

ΣERIM

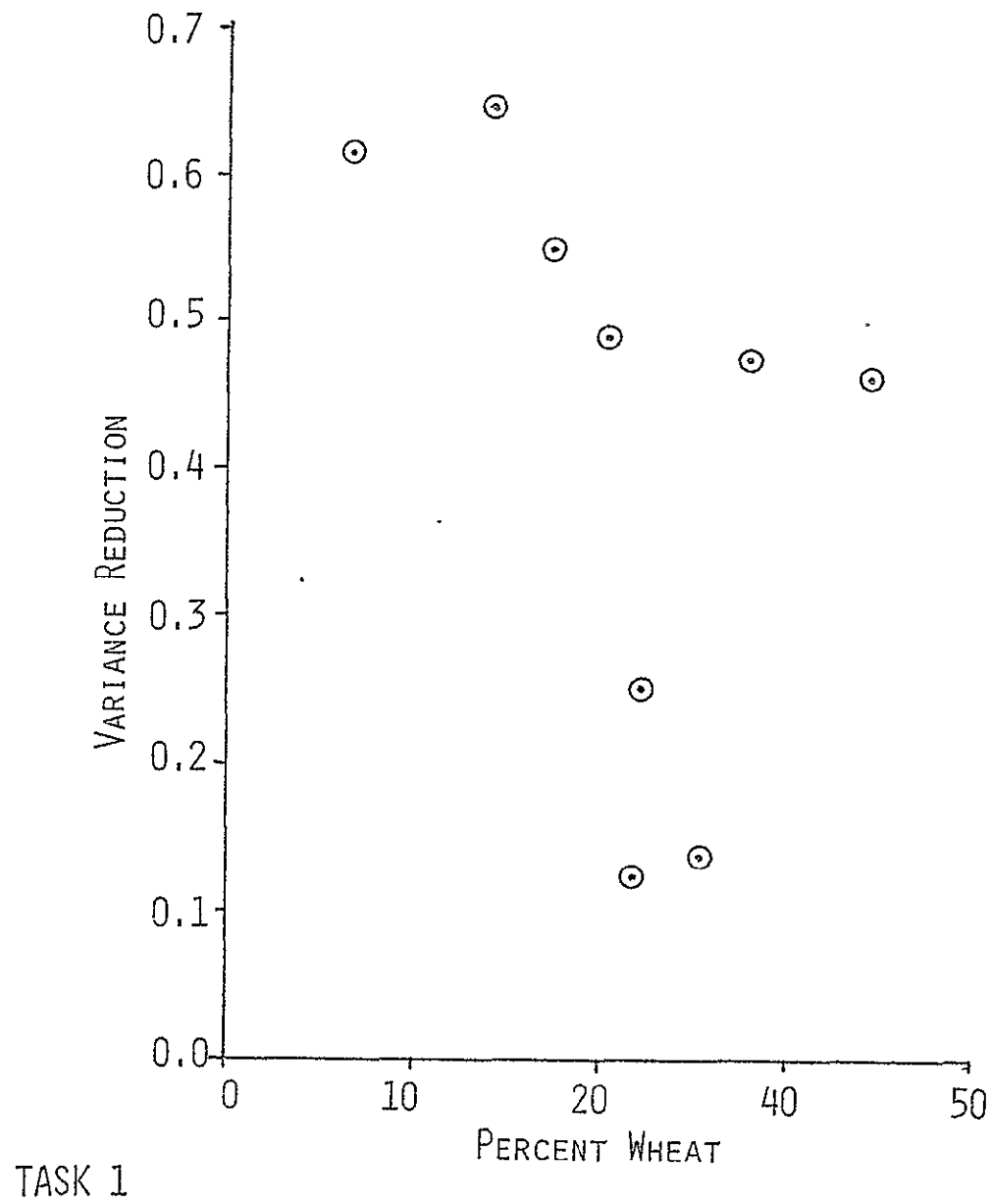


VARIANCE REDUCTION FACTOR AS A FUNCTION OF NUMBER OF BCLUSTERS FOR SINGLE SEGMENTS. VARIANCE REDUCTION FACTOR DUE TO BLOBBING IS SHOWN AS A LOWER BOUND. SLIDE 1 OF 2.

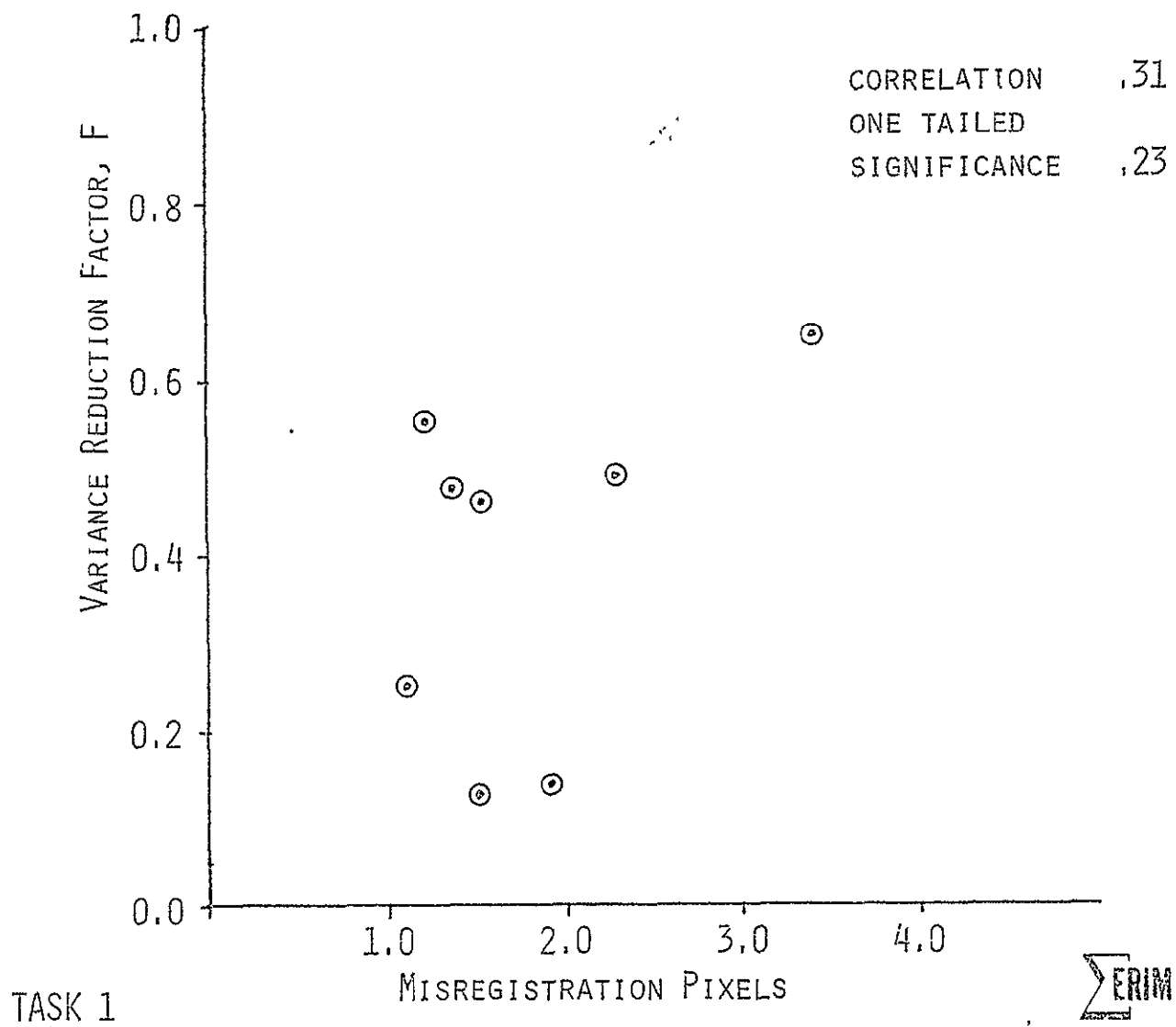
TASK 1

ΣERIM

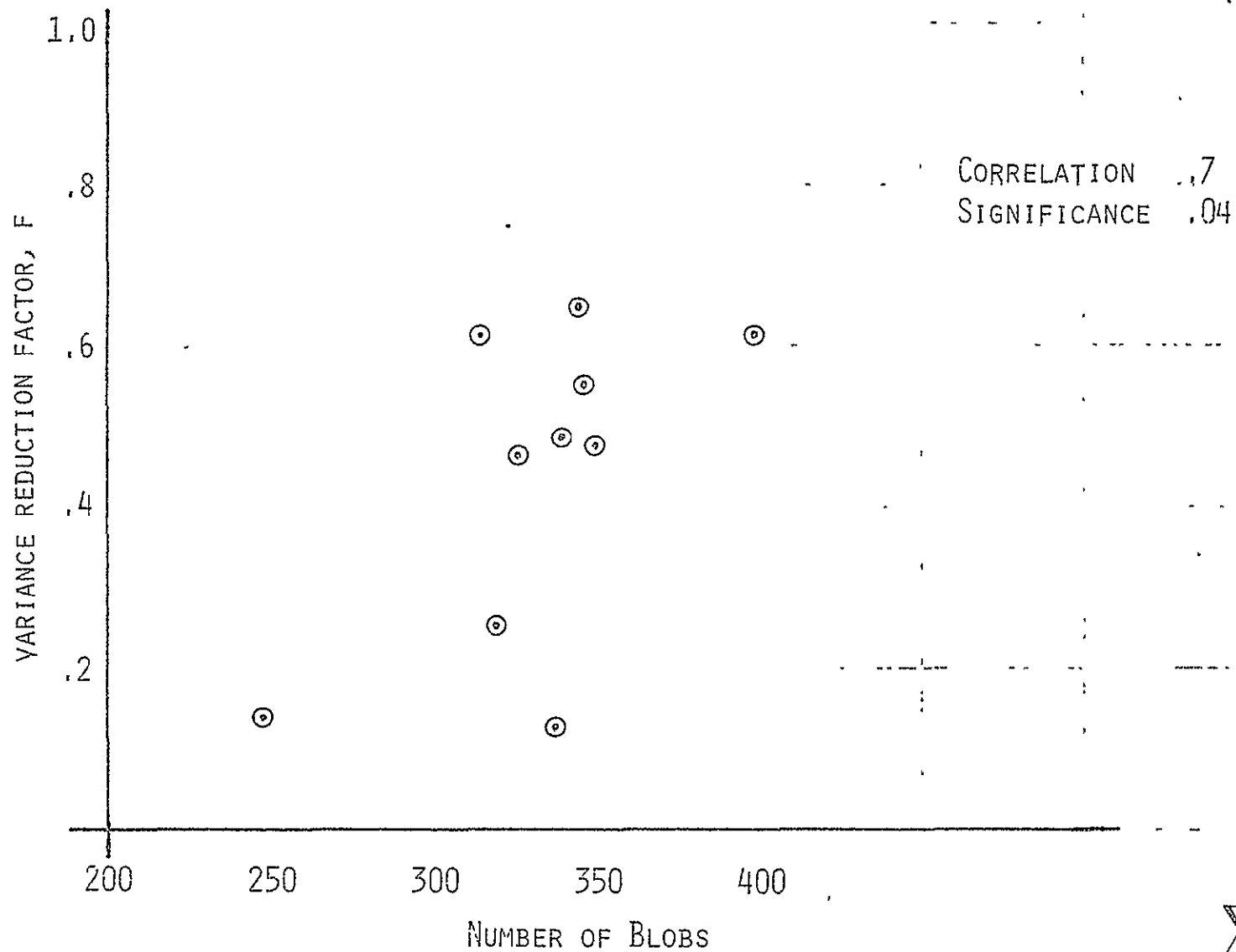
VARIANCE REDUCTION FACTOR FOR SINGLE SEGMENTS AS A FUNCTION OF THE PERCENT WHEAT



VARIANCE REDUCTION FACTOR, F, VS MAXIMUM MISREGISTRATION
ESTIMATE (PIXELS)



VARIANCE REDUCTION FACTOR, F, vs NUMBER OF BLOBS WHICH SURVIVE AFTER STRIPPING



ORIGINAL PAGE IS
OF POOR QUALITY

TASK 1

Σ ERIM

TASK 2

EVALUATION OF PARTITIONS FOR SIGNATURE EXTENSION

(R. C. Cicone)*

2.1 INTRODUCTION

The sampling and classification strategy of the Large Area Crop Inventory Experiment (LACIE) entails employing local signature training to determine wheat proportion estimates within 5x6-mile sample segments. Wheat proportions are then aggregated within designated strata. Multi-segment signature extension is philosophically founded on the premise that representative training information may be determined using non-local procedures at an additional savings in cost and reduction in the variance of the estimate.

Task 2 is concerned with addressing the key issues found in Table 2.1 that pertain to non-local training techniques.

2.2 OBJECTIVE

The objective of Task 2 is to test and evaluate techniques and procedures which embody the signature extension approach to large area crop inventories using Landsat data.

2.3 APPROACH

The approach adopted to address the objective of Task 2 is outlined in Table 2.2

2.4 SUMMARY OF PROGRESS DURING SECOND QUARTER

Table 2.3 reviews progress of Task 2 during the first quarter of the contract. Progress during the second quarter is outlined in Table 2.4.

* T. Wessling and J. Stinson contributed to the work reported.

2.5 DETAILS OF PROGRESS

Efforts this quarter have concentrated on a parametric evaluation of multisegment signature extension in North Dakota using a technique called preclassification. Figure 2.1 diagrams the effort carried out. Results will be described for cases employing Biowindows 1 and 2 in this spring wheat growing state, uncorrected and sun angle/haze corrected data, without partitioning of segments.

The preclassification technique employed is briefly described in Table 2.5. A more detailed description of this technique is available in Reference 1. The North Dakota data set used in this analysis is illustrated in Figure 2.2.

Of primary interest to this task is the role of partitioning in a multisegment signature extension environment. Figures 2.3(a), (b), and (c) illustrate a partitioning of North Dakota according to ten-year averages of precipitation and degree day parameters [2]. The analysis initiated in this quarter is concerned with techniques used to make use of static partitions. Several approaches are outlined in Table 2.6.

Figures 2.4 and 2.5 illustrate analysis conducted to date in an unstratified multisegment signature extension environment.

The ordinate axis of Figure 2.4 represents classification accuracies; the ordinate axis of Figure 2.5 represents bias in 209 dot grain proportion estimates. The abscissa specifies the number of segments used for training. Computations were carried out using different combinations of training segments drawn randomly without replacement. The graphs contain the average result with one standard deviation bar drawn about the average.

Figures 2.4(a) and 2.4(b) illustrate field center classification accuracy for grain and non-grain classes. Figure 2.4(a) was derived from uncorrected Landsat data. Figure 2.4(b) illustrates an improved grain classification accuracy using haze and sun angle corrected data [3]. Note the overall poor average classification accuracy achieved for grains without stratification of the segments.

Figures 2.5(a) and 2.5(b) illustrate the bias in 209 dot grain proportion estimate. The variance indicated is that due to the random draw of training segments, not that due to segment-to-segment variance in accuracy. The uncorrected case (Figure 2.5(a)) illustrates a tendency to underestimate grain by about 1% in the first case to over 7% in the last. The bias in the XSTAR corrected data is more stable across all cases beginning at about 1.7% to just over 3%. In any case the classifier displayed high variance segment-to-segment accuracies (not illustrated).

Corrected data seems to be establishing a pattern of improved results. This is supported also by the analysis reported last quarter using Kansas field mean data. It is yet to be seen what improvements can be made in a stratified environment. It is clear, however, that in an unstratified environment, multisegment signature extension is at a loss in coping with overlap of spectral distribution of different real classes that is most likely attributable to a number of ancillary conditions.

2.6 CONCLUSIONS AND RECOMMENDATIONS

Tables 2.7 and 2.8 contain conclusions and recommendations drawn based on the analyses conducted through the first two quarters of this task.

2.7 PLANS

Table 2.9 lists the activities to be conducted through the remainder of this contract year.

REFERENCES

1. Cicone, R. C., J. L. Stinson, and R. J. Balon, "An Evaluation of the Signature Extension Approach to Large Area Crop Inventories Utilizing Space Image Data", ERIM 122700-33-F, Environmental Research Institute of Michigan, Ann Arbor, Michigan, November 1977.
2. Colwell, R. N., C. M. Hay, R. W. Thomas, and A. S. Benson, "Development of Techniques for Producing Static Strata Maps and Development of Photointerpretation Methods Based on Multitemporal Landsat Data", Interim Report for NASA/JSC, University of California, Berkeley, 15 December 1976.
3. Lambeck, P. F., "Signature Extension Preprocessing for Landsat MSS Data", ERIM 122700-32-F, Environmental Research Institute of Michigan, Ann Arbor, Michigan, November 1977.

TABLE 2.1
KEY ISSUES

- MULTISEGMENT AREA ESTIMATION PROCEDURES
- SIGNATURE EXTENSION
- USE OF STATIC PARTITION BOUNDARIES
- DATA NORMALIZATION AND PREPROCESSING

TASK 2

TABLE 2.2

APPROACH

- TEST DESIGN AND IMPLEMENTATION CONSIDERATIONS
 - EXPERIMENT DESIGN
 - PROGRAM DEVELOPMENT
 - DATA PREPARATION
- PARAMETRIC EVALUATION USING PRECLASSIFICATION
 - TRAINING GAIN
 - HAZE CORRECTION
 - STATIC PARTITION BOUNDARIES
 - TRAINING SEGMENT SELECTION
- PROCEDURE EVALUATION
 - LOCAL PROCEDURE B
 - PROCEDURE B
 - MULTISEGMENT PROCEDURE 1

TABLE 2.3

REVIEW OF PREVIOUS QUARTER'S PROGRESS

- EVALUATED MULTISEGMENT SIGNATURE EXTENSION
 - USING 1975-76 KANSAS FIELD MEANS DATA BASE
 - PRECLASSIFICATION TECHNIQUE
- XSTAR HAZE AND SUN ANGLE CORRECTION RESULTED IN SUBSTANTIAL IMPROVEMENT IN FIELD CENTER CLASSIFICATION ACCURACY
- THE USE OF UCB STATIC DEGREE DAYS AND PRECIPITATION STRATA SIGNIFICANTLY IMPROVED RESULTS, AT A COST IN TRAINING GAIN
- THE USE OF A GREEN DEVELOPMENT INDICATOR AS A SEGMENT MATCHING CRITERION IN BIOWINDOW 1 RESULTED IN IMPROVED CLASSIFICATION ACCURACY

TABLE 2.4

SUMMARY OF PROGRESS

- PROGRAM DEVELOPMENT

- PRECLAS, TASCAP, STRIP, COMPRS, GNDTRU, ILEC

- PREPARED NORTH DAKOTA DATA SET

- 16 LACIE PHASE III BLIND SITES
- 4 BIOWINDOWS
- JSC GROUND TRUTH

- EVALUATION OF MULTISEGMENT SIGNATURE EXTENSION USING PRECLASSIFICATION

- PROPORTION ESTIMATION + FIELD CENTER ACCURACY
- 2 AND 3 BIOWINDOWS (15 AND 11 SEGMENTS)
- CORRECTED AND RAW DATA
- UCB PARTITIONS

TABLE 2.5

PRECLASSIFICATION

- A SUM OF LIKELIHOODS CLASSIFIER
 - COST-EFFECTIVE MEANS OF TESTING
 - CAN EXAMINE SIGNIFICANCE OF A PRIORI WEIGHTING FACTORS
 - FACILITATES OPERATIONAL USE OF PARTITIONS WITH MULTISEGMENT PROCEDURES
- EACH SEGMENT CLASSIFIED BY EACH TRAINING SEGMENT SEPARATELY USING SUM OF LIKELIHOODS, QUADRATIC CLASSIFIER
- DECISION MADE BY COMPARING APPROPRIATELY WEIGHTED SUMS OF WHEAT AND OTHER LIKELIHOODS FROM DIFFERENT TRAINING SEGMENTS

TABLE 2.6

USE OF PARTITIONS

- UCB DEGREE DAYS, PRECIPITATION
- STRATEGIES CONSIDERED
 - TREAT BOUNDARIES AS STATIC
 - WEIGHT TRAINING FROM SAME STRATA MORE HEAVILY
 - LIMIT TRAINING TO SAME OR ADJACENT STRATA
 - USE PARTITION VARIABLES AS DISTANCE MEASURES FOR WEIGHTING TRAINING
 - INSURE EACH PARTITION IS REPRESENTED BY A TRAINING SEGMENT
- USE FUNCTION OF DISTANCE IN MILES BETWEEN SEGMENTS FOR WEIGHTING SIGNATURES

TABLE 2.7
CONCLUSIONS

- UNSTRATIFIED SIGNATURE EXTENSION IS AFFECTED BY UNDESIRABLE
"SEGMENT-TO-SEGMENT OVERLAP BETWEEN GRAIN AND OTHER SIGNATURES,"
WITH CLASSIFICATION DOMINATED BY MORE PREVALENT OTHER CLASSES.
- PREPROCESSING TECHNIQUES APPEAR TO BE UN-DOING THE PARTIAL OVERLAP
OF OTHERWISE SEPARABLE GRAIN AND OTHER PIXELS WHICH IS DUE TO
SCENE-TO-SCENE HAZE AND SUN ANGLE VARIABILITY,

TABLE 2.8

RECOMMENDATIONS

- IN TWO MULTISEGMENT TESTS CONDUCTED EVEN THE BEST RESULTS IN AN UNSTRATIFIED ENVIRONMENT WERE POOR. AN UNDERSTANDING OF HOW TO USE STRATIFICATION TO IMPROVE THESE RESULTS MUST FIRST BE ATTAINED BEFORE USING MULTISEGMENT SIGNATURE EXTENSION OPERATIONALLY.
- HAZE AND SUN ANGLE PREPROCESSING SHOULD BE INCORPORATED INTO ANY PROCEDURE EMPLOYING SIGNATURE EXTENSION TECHNIQUES.

TABLE 2.9

PLANS

- EXAMINE ROLE OF STATIC STRATIFICATION
- DOCUMENT RESULTS
- TEST LOCAL PROCEDURE B AND COMPARE TO PROCEDURE 1 APPROACH
- EXAMINE USES OF AMOEBA IN PROCEDURE B
- TEST MULTISEGMENT PROCEDURE B

FIGURE 2.1

EVALUATION OF MULTISEGMENT SIGNATURE EXTENSION
PROCEDURE

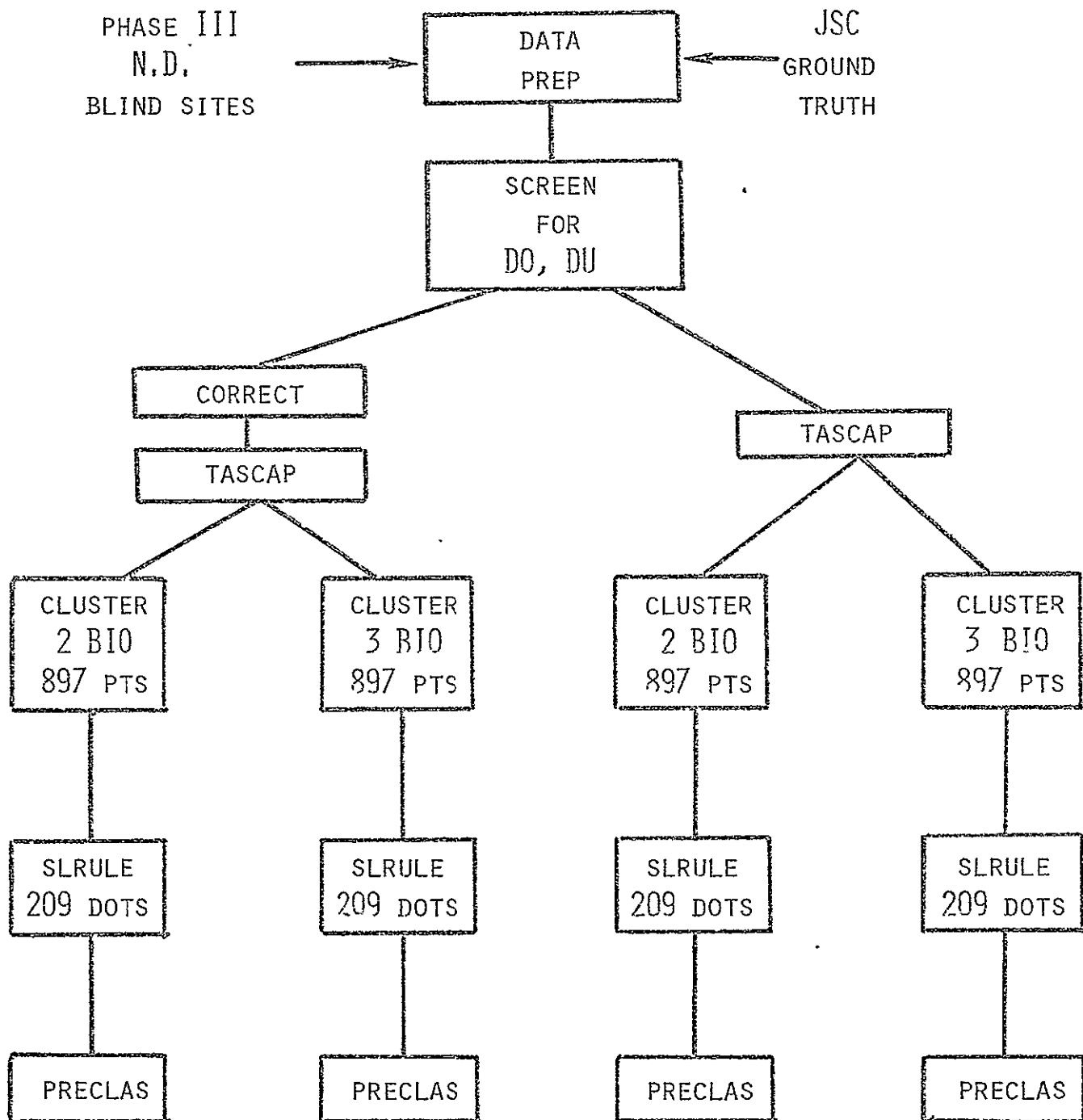
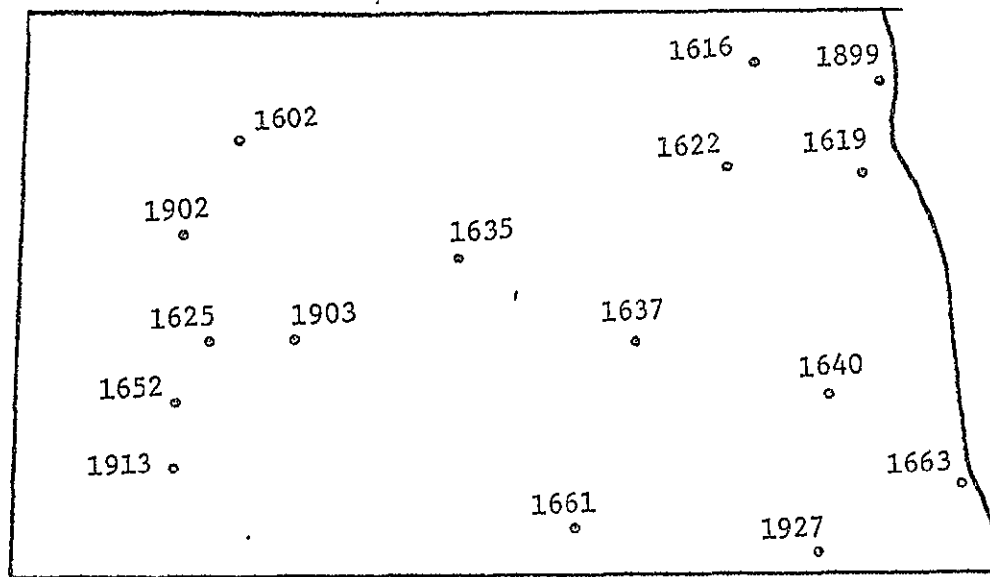


FIGURE 2.2

NORTH DAKOTA DATA SET

2-15

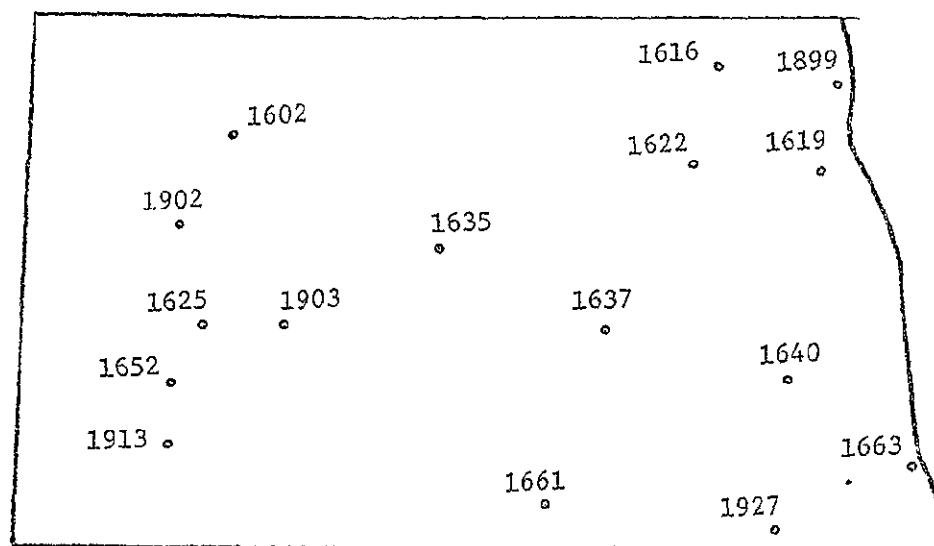


BIOWINDOWS

	1	2	3	4
1602	✓	✓	✓	✓
1616	✓	✓		
1619	✓	✓		
1622	✓	✓		
1625	✓	✓	✓	✓
1635	✓	✓	✓	✓
1637	✓	✓	✓	✓
1640	✓	✓	✓	✓
1652	✓	✓	✓	✓
1661	✓	✓		
1663	✓	✓	✓	✓
1899	✓	✓	✓	
1902	✓		✓	
1903	✓	✓	✓	✓
1913	✓	✓	✓	✓
1927	✓	✓	✓	✓

FIGURE 2.3(a)

UCB STATIC STRATIFICATION OF NORTH DAKOTA



2-16

TASK 2

Σ ERIM

FIGURE 2.3(b)

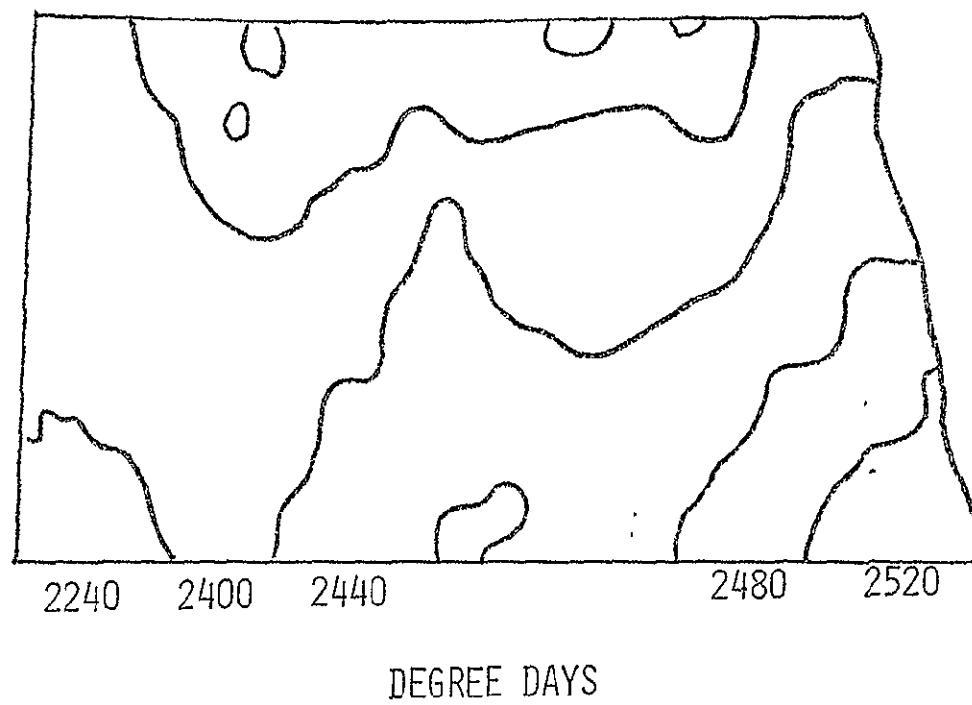
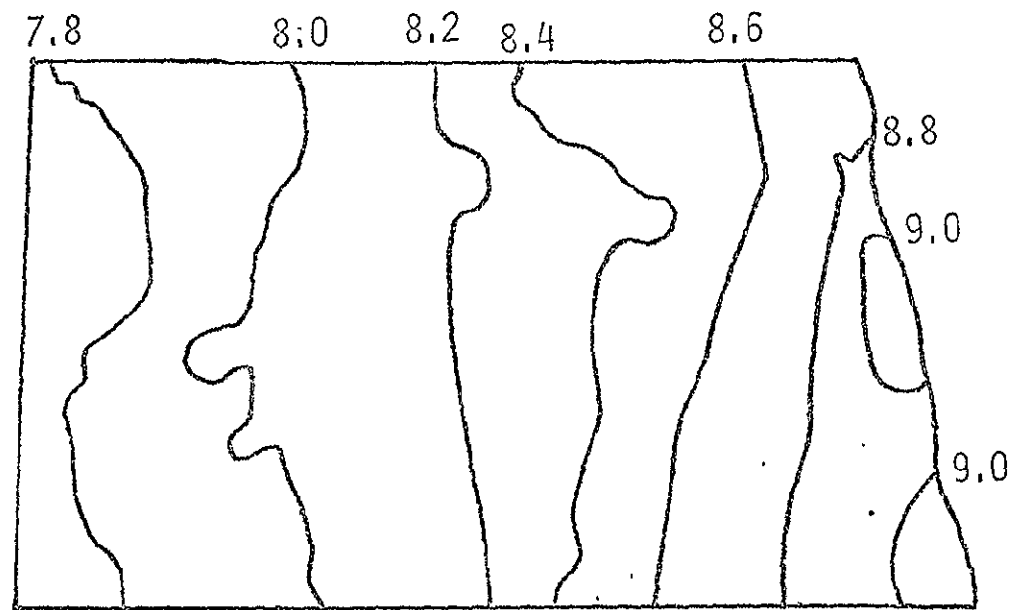


FIGURE 2.3(c)



PRECIPITATION

MULTISEGMENT SIGNATURE EXTENSION

15 N.D. PHASE III BLIND SITES

BIOWINDOWS-1 & 2

RAW DATA

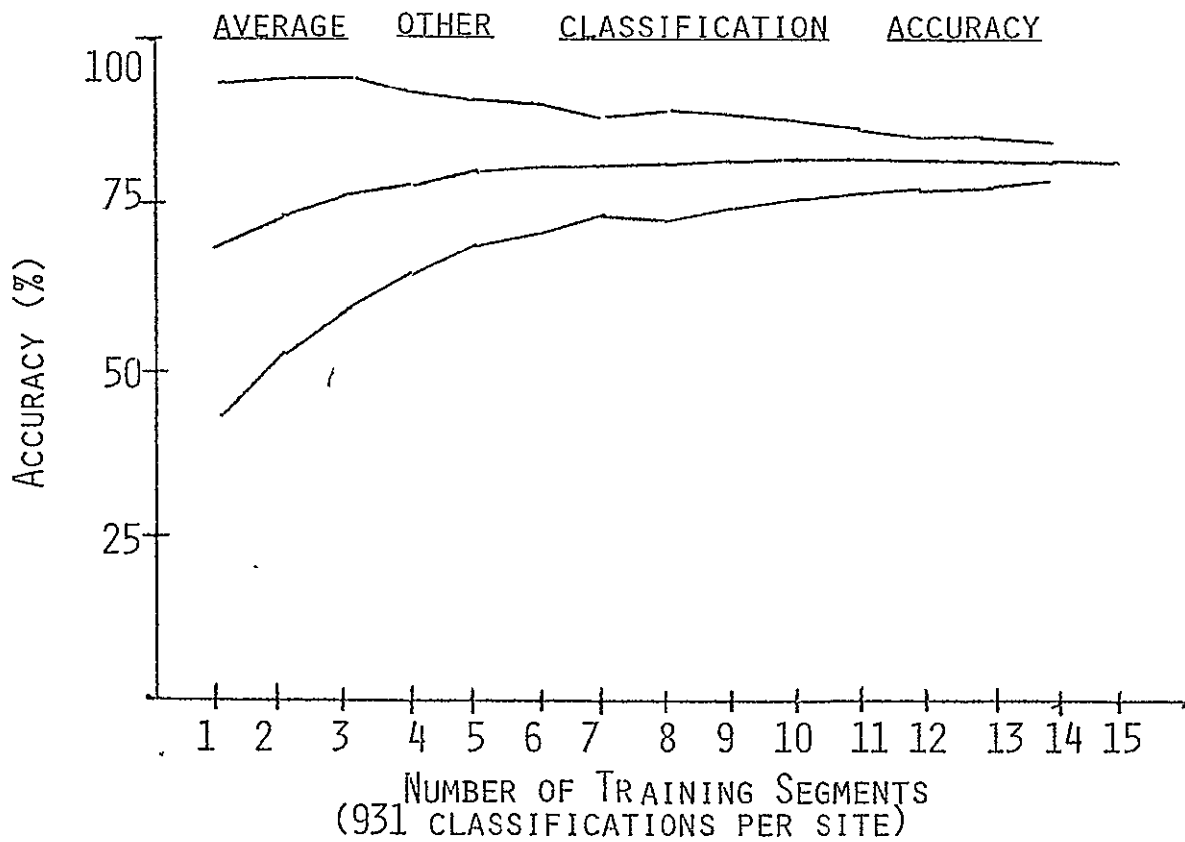
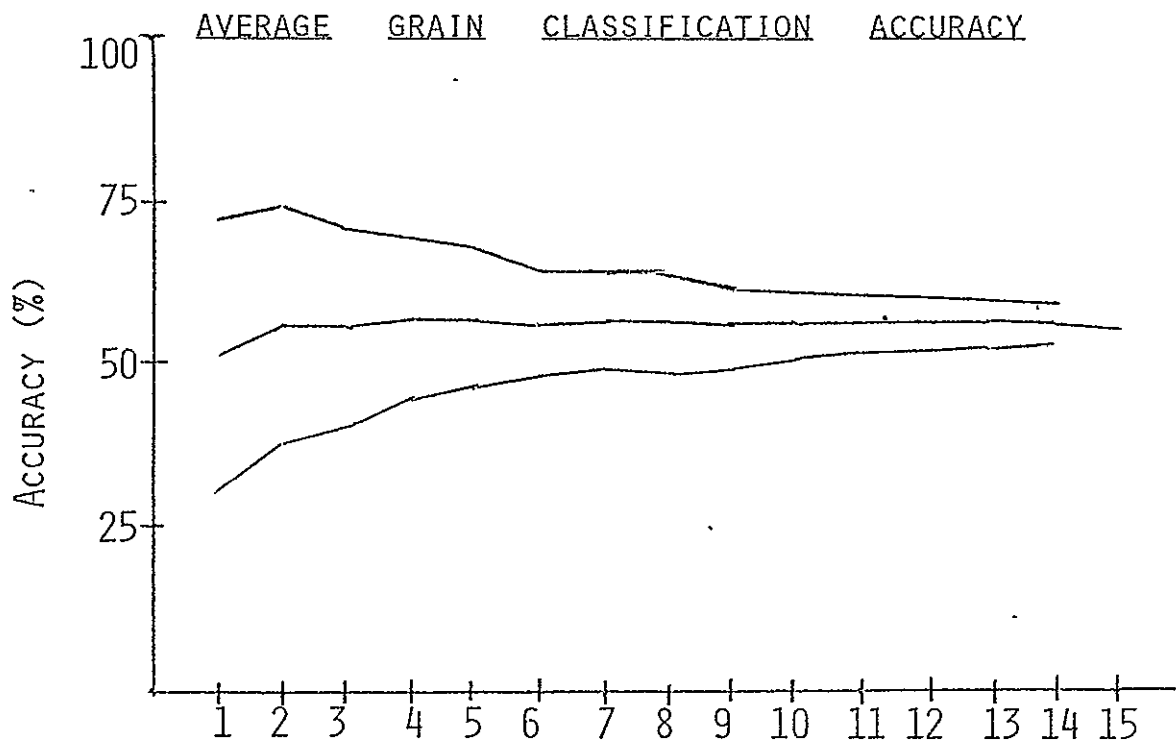


FIGURE 2.4(b)

X STAR DATA

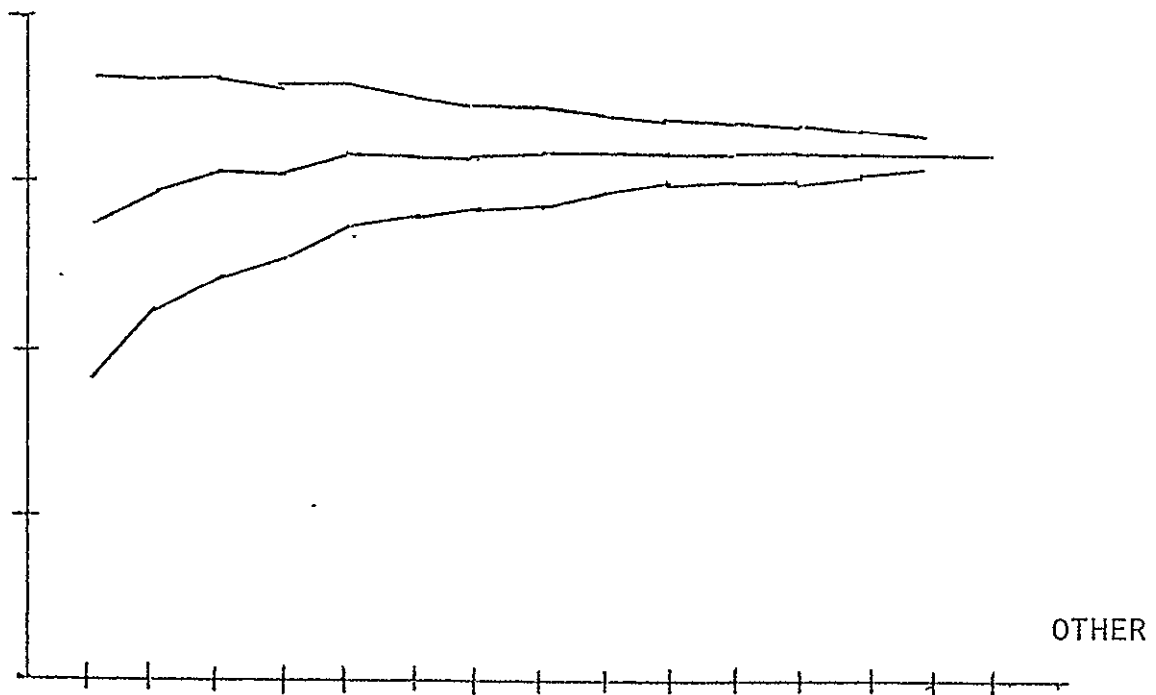
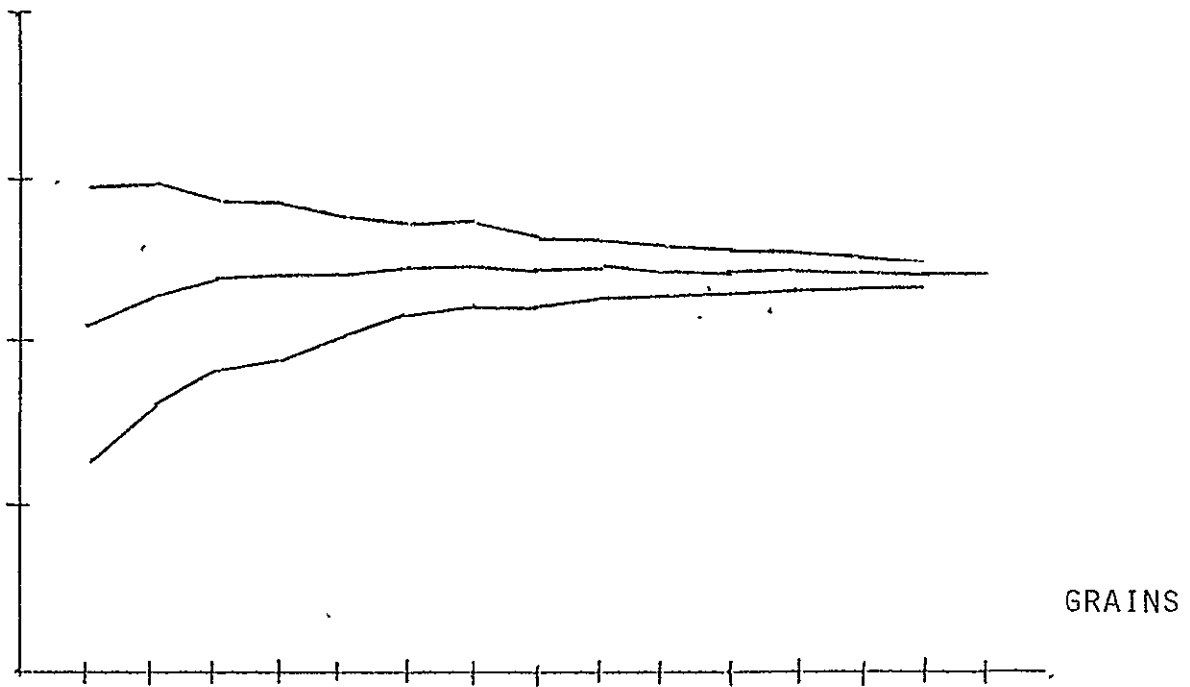


FIGURE 2.5(a)

MULTISEGMENT SIGNATURE EXTENSION

15 N.D. PHASE III BLIND SITES

BIOWINDOWS 1 & 2

209 DOT PROPORTION ESTIMATION ACCURACY

RAW DATA

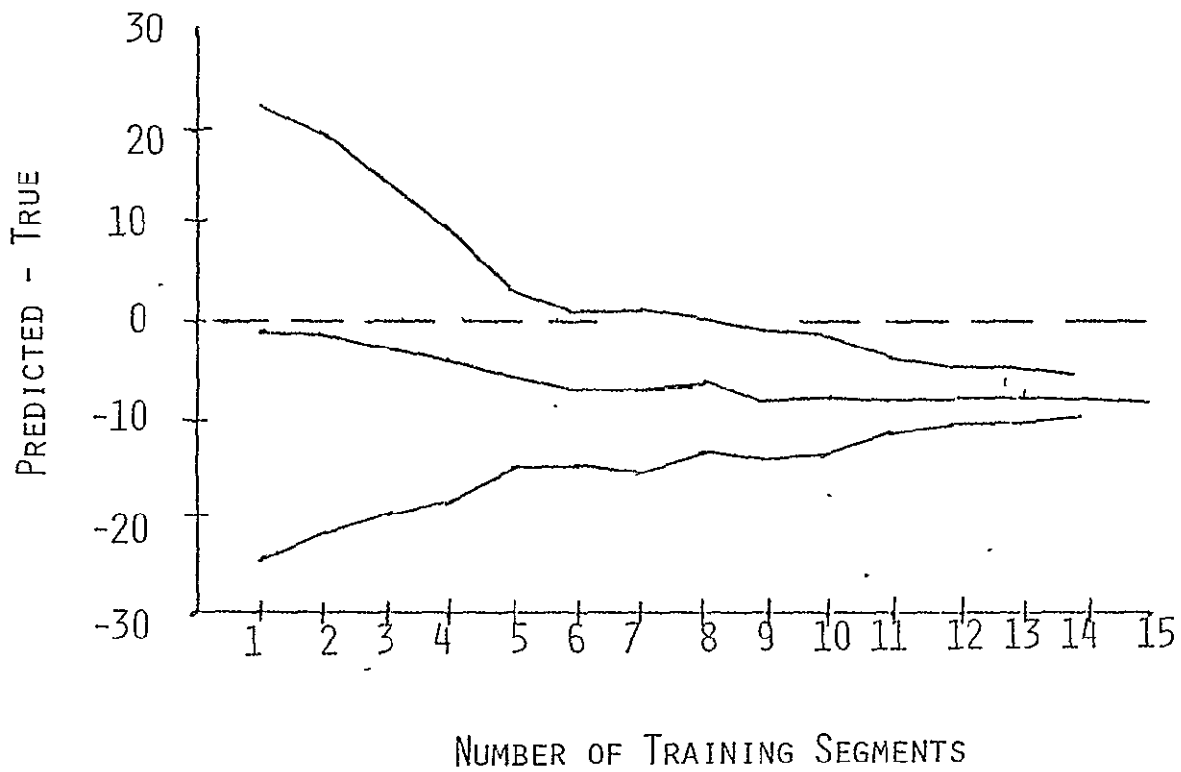
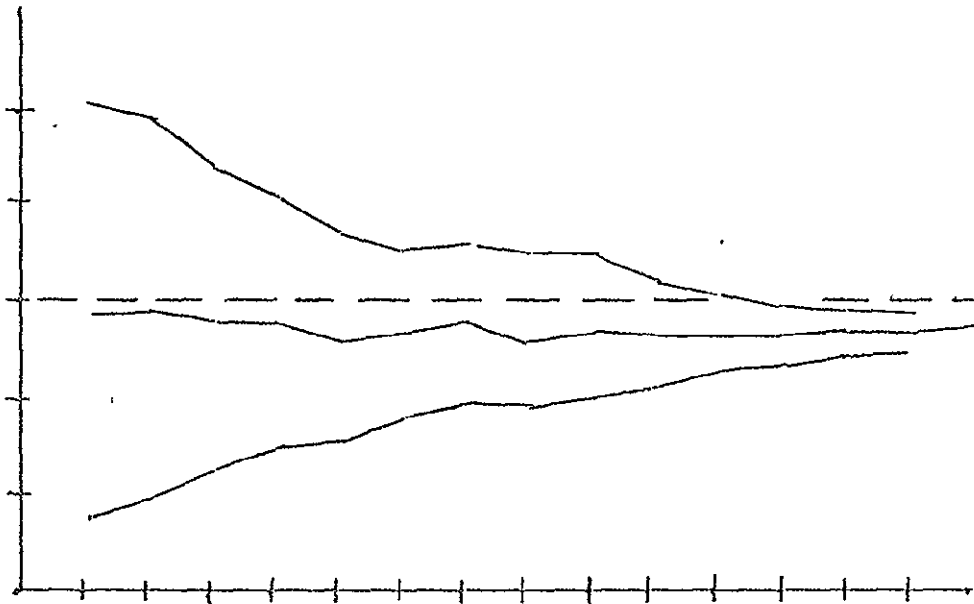


FIGURE 2.5(b)

XSTAR DATA

209 DOT ESTIMATE





FORMERLY WILLOW RUN LABORATORIES THE UNIVERSITY OF MICHIGAN

TASK 4

SPECTRAL SEPARABILITY OF SPRING WHEAT FROM OTHER SMALL GRAINS

(W.A. MALILA AND J.M. GLEASON*)

4.1 INTRODUCTION

The problem of distinguishing between spring wheat and other small grains is a problem of interest to LACIE and similar agricultural crop inventories. A prior study at ERIM** gave indications of separability of spring wheat and barley under certain conditions in Landsat multispectral data from several Phase 2 LACIE Blind Sites in North Dakota. Last quarter, we reported confirmation and explanation of many observed characteristics in the Landsat signatures of small grains, through analysis of 1975-76 LACIE Field Measurement data and USDA crop reports and crop production statistics. Also several implications for discrimination procedures were drawn from analysis of the results.

The objective of this Task is to develop a spectral classification method for discriminating spring wheat from other small grains, using Landsat data.

4.2 APPROACH

Table 4.1 summarizes the approach taken toward meeting the Task objective.

4.3 SUMMARY OF PROGRESS DURING THE SECOND QUARTER

During the reporting period, two data sets from the 1976-77 growing season were prepared and analyzed. First were LACIE field measurements of the reflectance of spring small grains, made in Williston, N. Dakota, using an Exotech Model 100 (Landsat-band)

* W. Malila served as Task leader; E.P. Crist also contributed to the reported work.

** Malila, William A. and James M. Gleason, "Investigations of Spectral Separability of Small Grains, Early Season Wheat Detection, and Multi-crop Inventory Planning, "ERIM 122700-34-F, Environmental Research Institute of Michigan, Ann Arbor, Michigan, November 1977.



radiometer. Second were LACIE Phase 3 Blind Site data from North and South Dakota, Montana, and Minnesota. Substantial data preparation activities were conducted to produce a comprehensive Phase 3 data set - included were screening, corrections for haze and sun angle, sampling on a 5x5 grid (897 points), and merging with crop labels extracted from JSC ground truth tapes.

Trends observed in LACIE Phase 2 data and corresponding field measurements were also found in these more recent data sets. Unitemporal correct classification percentages for spring wheat vs. barley were in the 80's given complete training on individual segments and 76% for a multisegment example. Spring wheat and barley were most separable in the growth stage when they turn color from green to yellow-green or yellow (the soft dough stage). Barley tends to turn color sooner than wheat and some of it has a distinctive brightness after turning is complete. Tasseled - cap brightness again was the key feature for discrimination; together with greenness it captured most data variation and spectral differences. A parameter indicating crop maturity would appear promising for exploring differences in crop calendar, where coupled with crop calendar estimates.

The best linear discriminant tended to overestimate the ratio of spring wheat to barley, even given complete training on the small grain observations.

Finally, from this our first analysis of extensive Landsat data on oats, we found that the spectral separation of spring wheat from oats to be substantially more difficult than from barley.



4.4 DETAILS OF SECOND QUARTER ACTIVITIES

4.4.1 ANALYSIS OF FIELD MEASUREMENT DATA

The data set analyzed is described in Table 4.2. It represents a portion of the LACIE Field Measurements made in North Dakota during the 1976-77 growing season. The instrument has filters which cover the spectral bands of Landsat. Linear discriminant analysis was performed on these measurements as a function of observation date. The results presented in Table 4.3 represent an upper bound on separability that might be achieved. Note that separability was best after heading. The next part of our analysis was to examine LACIE Phase 3 Landsat data directly.

4.4.2 ANALYSIS OF LACIE PHASE 3 LANDSAT DATA

An extensive data set was assembled by processing data provided by NASA/JSC. Landsat data from 28 segments in four states (See Figure 4.1) were prepared in a manner described in Table 4.4. Preprocessing included corrections for atmospheric haze and sun angle and a systematic sampling of the pixels. The 897 points selected include the 209 dots of Procedure 1 as a subset.

Twelve segments with substantial numbers of barley pixels were analyzed for (linear discriminant) separability of labeled spring wheat and barley using Landsat spectral data. The results presented in Table 4.5 for individual segments by acquisition date show clearly that in 1977 the July 12-13 time period (Julian Day 193-194) was the best date for separability. Most correct classification values are in the 80% range. Each number in the table is the conditional average correct classification, i.e., the average of the spring wheat and barley correct classification values. It was noted that the spring wheat value generally was larger than the barley value.

We wished to rank order the tasseled-cap features in their contribution to separating spring wheat and barley. Seven segments (excluding Montana 1929) had acquisitions on the best date. First, their features were rank ordered on an individual segment basis, with results as shown in Table 4.6. Next, haze-corrected data values were pooled, a composite multisegment linear decision rule was established, and the tasseled cap features were ranked. Brightness was most important in both the composite and most individual segments. The second-ranked feature varied from segment to segment, but greenness was selected for the composite data set.

The location of the multisegment decision line in Greenness-Brightness space is shown in Figure 4.2 with respect to all spring wheat points and in Figure 4.3 with respect to all barley points. The classification accuracies achieved by this line (four-dimensional version) are presented on the right half of Table 4.6. Overall, wheat was 85% correct, barley 68%, and the average 76% correct. Values for individual segments also are given. Note the atypical results obtained for segment 1699 and the explanation in the footnote. This result highlights the need for a good crop calendar estimate to help establish decision lines between spring wheat and barley.

In addition, we examined the spectral separability of oats and spring wheat in Twelve segments. (Our LACIE Phase 2 data set did not have sufficient oats for earlier comparisons.) The results presented in Table 4.7 indicate both poorer separability than spring wheat vs. barley and no clear best date for separation.

Finally, we performed multitemporal clustering of spring wheat and barley points for one segment (1663) to better examine temporal trajectories and their variability. Plots for four wheat clusters and two barley clusters are presented in Figures 4.4 and 4.5, respectively.



4.5 RECOMMENDATION

It is recommended that crop growth stage information be included in the information acquired as part of the periodic observations of fields in blind sites.

4.6 PLANS

Plans for the third quarter are presented in Table 4.8.

TABLE 4.1

APPROACH

- GAIN UNDERSTANDING OF POTENTIAL DISCRIMINATION FEATURES, SUCH AS DIFFERING SPECTRAL CHARACTERISTICS VS. CROP CALENDAR AND DIFFERING CROP CALENDARS
 - BY USING AVAILABLE LANDSAT DATA
 - SUPPLEMENTED BY OTHER DATA
 - • LACIE FIELD MEASUREMENT DATA
 - • USDA AGRICULTURAL STATISTICS
- ESTABLISH, TEST AND EVALUATE ONE OR MORE CLASSIFICATION PROCEDURES
- IF RESULTS SUFFICIENTLY PROMISING, ESTABLISH SENSITIVITY OF CLASSIFICATION PERFORMANCE TO ACQUISITION SEQUENCE AND ACCURACY OF TRAINING
- BEGIN CONSIDERATION OF ADVANCED PROCEDURES, E.G., YEAR-TO-YEAR OR USE OF THEMATIC MAPPER BANDS

TABLE 4.2

SMALL GRAINS REFLECTANCE DATA

1977 LACIE FIELD MEASUREMENTS IN WILLIAMS COUNTY, NORTH DAKOTA;
WITH EXOTECH MODEL 100 RADIOMETER

• 32 PLOTS

4 CROPS: HARD RED SPRING WHEAT
 DURUM SPRING WHEAT
 BARLEY
 OATS

2 VARIETIES EACH CROP

2 SOIL MOISTURE LEVELS: FALLOW IN 1976
 WHEAT IN 1976

2 REPLICATES EACH COMBINATION

• 6 MEASUREMENT DATES

	<u>JULIAN</u>	<u>SPRING WHEAT GROWTH STAGE</u>	<u>KEY TO GROWTH STAGE</u>
6/18	169	2	2 TILLERING
6/23	174		3 STEM EXTENSION
7/03	184	3,4	4 HEADING AND FLOWERING
7/14	195	4	5 RIPENING
7/20	201	5,4	
8/05	217	5	

TASK 4

TABLE 4.3

LINEAR DISCRIMINATION RESULTS ON 1977 RADIOMETER REFLECTANCE DATA
 SPRING WHEAT (HARD RED AND DURUM) VS. OTHER SMALL GRAINS (BARLEY AND OATS)

CLASSIFICATION PERCENTAGE

JUN 18 (DAY 169)

TRUE CLASS

DECISION CLASS	<u>TRUE CLASS</u>		AVG. CORRECT
	SW	OSG	
SW	55	23	66
OSG	45	77	
No. OBSERVATIONS		29 30	

JUN 23 (DAY 174)

TRUE CLASS

DECISION CLASS	<u>TRUE CLASS</u>		AVG. CORRECT
	SW	OSG	
SW	75	25	75
OSG	25	75	
No. OBSERVATIONS		16 16	

JUL 3 (DAY 184)

TRUE CLASS

DECISION CLASS	<u>TRUE CLASS</u>		AVG. CORRECT
	SW	OSG	
SW	81	31	75
OSG	19	69	
No. OBSERVATIONS		16 16	

JUL 14 (DAY 195)

TRUE CLASS

DECISION CLASS	<u>TRUE CLASS</u>		AVG. CORRECT
	SW	OSG	
SW	83	13	85
OSG	17	87	
No. OBSERVATIONS		58 60	

JUL 20 (DAY 201)

TRUE CLASS

DECISION CLASS	<u>TRUE CLASS</u>		AVG. CORRECT
	SW	OSG	
SW	94	6	94
OSG	6	94	
No. OBSERVATIONS		16 16	

AUG 5 (DAY 217)

TRUE CLASS

DECISION CLASS	<u>TRUE CLASS</u>		AVG. CORRECT
	SW	OSG	
SW	96	3	96*
OSG	4	97	
No. OBSERVATIONS		76 30	

TASK 4

*3/4 OF OATS AND 1/2 BARLEY HARVESTED BY AUG. 5.



TABLE 4.4

SPRING SMALL GRAINS DATA SET PREPARED FROM LACIE PHASE 3 BLIND SITES (1976-77)

- 28 SEGMENTS FROM FOUR STATES: 16 NORTH DAKOTA
 6 SOUTH DAKOTA
 4 MONTANA
 2 MINNESOTA

- PREPROCESSED LANDSAT DATA
 - SCREENED VISUALLY AND DIGITALLY TO FLAG CLOUDS, SHADOWS,
AND BAD DATA
 - ACQUISITIONS WERE SELECTED
 - CORRECTED FOR HAZE AND SUN ANGLE (XSTAR ALGORITHM)
 - TRANSFORMED TO TASSELLED-CAP FEATURES

- SAMPLED 897 POINTS (5x5 GRID, INCLUDES 209 DOTS AS SUBSET)

- MERGED WITH GROUND TRUTH DATA FROM JSC TAPES (LABELED AS EITHER
A SINGLE SPECIFIC COVER CODE OR A MIXTURE PIXEL)

- ASSOCIATED ANCILLARY DATA
 - LACIE CROP CALENDAR ESTIMATES
 - LATITUDE AND LONGITUDE

TABLE 4.5

SPRING WHEAT VS. BARLEY DISCRIMINATION

LANDSAT DATA FROM LACIE PHASE 3 BLIND SITES (FROM 897 POINTS)

COMPLETE TRAINING USING GROUND TRUTH:

LINEAR DECISION RULE.

		AVERAGE PERCENT CORRECT ON DATE:								
STATE	SEGMENT	No. PIXELS		MAY* 1-3	MAY 18-21	JUN 5-8	JUN 23-25	JUL 12-13	JUL 29-31	Aug 17-18
		SW	B							
MINN	1515	287	142			63	56	85**		
	1523	139	75	59	66	64	56		68	67
ND	1616	288	185	59	67	63				62
	1619	315	67	55	55	67	72		60	65
	1622	252	103	60	66	61	57			56
	1637	194	28	65	55	54		83		
	1640	224	103	60	69		64	78	78	69
	1663	239	93	60	64	73	69	77	65	56
	1899	308	224	52	61	65	62	80		
	1927	127	47	57	67	64	71	85		58
SD	1699	54	23		72	66	89	97		64
MONT	1929*	61	21	72	70		70	69	88	

* LANDSAT ACQUISITIONS OVER MONTANA WERE APPROXIMATELY ONE WEEK AFTER OTHER ACQUISITIONS.

** BOX DENOTES LARGEST VALUE FOR SEGMENT.

TABLE 4.6

RANKING OF TASSELLED CAP FEATURES FOR DISCRIMINATING BETWEEN
SPRING WHEAT AND BARLEY

JULY 12,13 ACQUISITIONS (1973)
(DAYS 193,194)

XSTAR-CORRECTED LANDSAT DATA

<u>SEGMENT</u>		<u>ORDER OF CHOICE</u>	<u>CROP CALENDAR</u>	<u>MULTISEGMENT RULE % CORRECT CLASSIFICATION</u>		
				<u>SP WHT</u>	<u>BARLEY</u>	<u>AVG.</u>
MINN	1515	B,G,Y,N	5.4	90	77	84
SD	1699	B,N,G,Y	6.0	0*	100	50
ND	1637	N,B,G,Y	5.3	92	54	73
	1640	G,B,Y,N	5.4	88	61	75
	1663	B,G,Y,N	5.3	83	74	78
	1899	B,G,N,Y	5.3	93	58	75
	1927	B,G,Y,N	5.4	81	81	81
ALL SEVEN		B,G,N,Y		85	68	76

* NOTE ADVANCED CROP CALENDAR FOR SEGMENT 1699 WHICH CAUSED SPRING WHEAT TO MOVE TO BARLEY SIDE OF DAY 193 DECISION LINE; ON DAY 176 WITH CROP CALENDAR 5.1, THE RESULT WAS SW = 96%, B = 48%, AND AVG. = 72%.

TABLE 4,7

SPRING WHEAT VS. OATS DISCRIMINATION

LANDSAT DATA FROM LACIE PHASE 3 BLIND SITES (FROM 897 POINTS)

COMPLETE TRAINING USING GROUND TRUTH.

LINEAR DECISION RULE.

STATE	SEGMENT	No. PIXELS		AVERAGE PERCENT CORRECT ON DATE:						
				MAY	MAY	JUN	JUN	JUL	JUL 29	AUG
		SW	B	1-5	18-23	5-8	23-28	12-16	AUG 4	17-21
4-12	MINN	1515	287 36			58	64	58		
		1523	139 58	55	56	49	60		57	56
	ND	1602	215 22	55	55		54	63	60	
		1637	194 22	67	61	68		71		
		1652	116 21	55	64		55	60		64
		1661	188 38	52		52				58
		1663	239 27	54	66	55	74	74	76	62
		1903	63 27	62			59	64		67
		1927	127 37	62	64	54	65	65	56	
	SD	1675	85 21	59	64		47			64
		1686	27 84		79			67	69	
		1699	54 66		72	68	77	73		64

TABLE 4.8

PLANS

- CONDUCT MORE EXTENSIVE STUDIES OF THE RELATIONSHIP BETWEEN PERIODIC GROUND OBSERVATIONS AND LANDSAT DATA FOR WHEAT IN PHASE 3 BLIND SITES
- CONTINUE ANALYSIS OF FIELD MEASUREMENT DATA
- DEVELOP AND TEST A SPRING WHEAT VS. BARLEY DISCRIMINATION PROCEDURE BASED ON A SPECTRAL CROP MATURITY INDEX

FIGURE 4.1

LOCATIONS OF LACIE PHASE 3 BLIND SITES ANALYZED

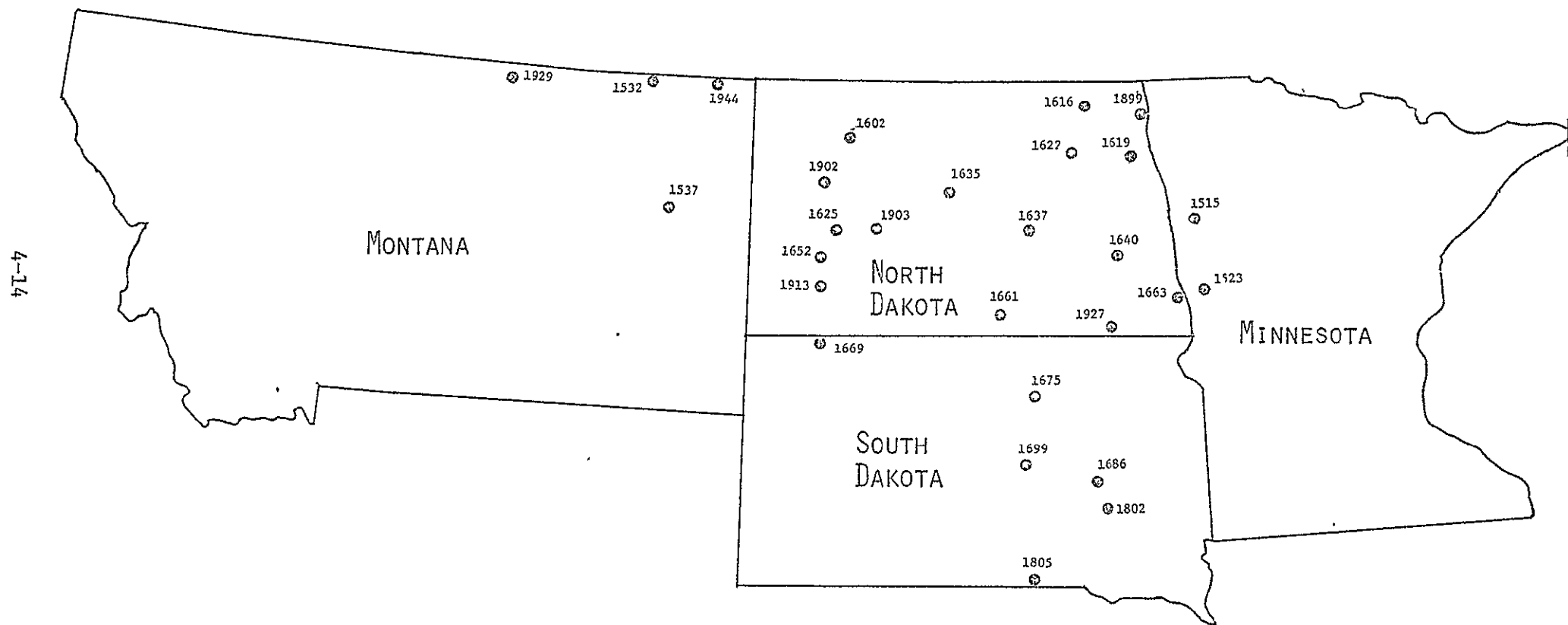


FIGURE 4.2
SCATTER OF SPRING WHEAT DATA

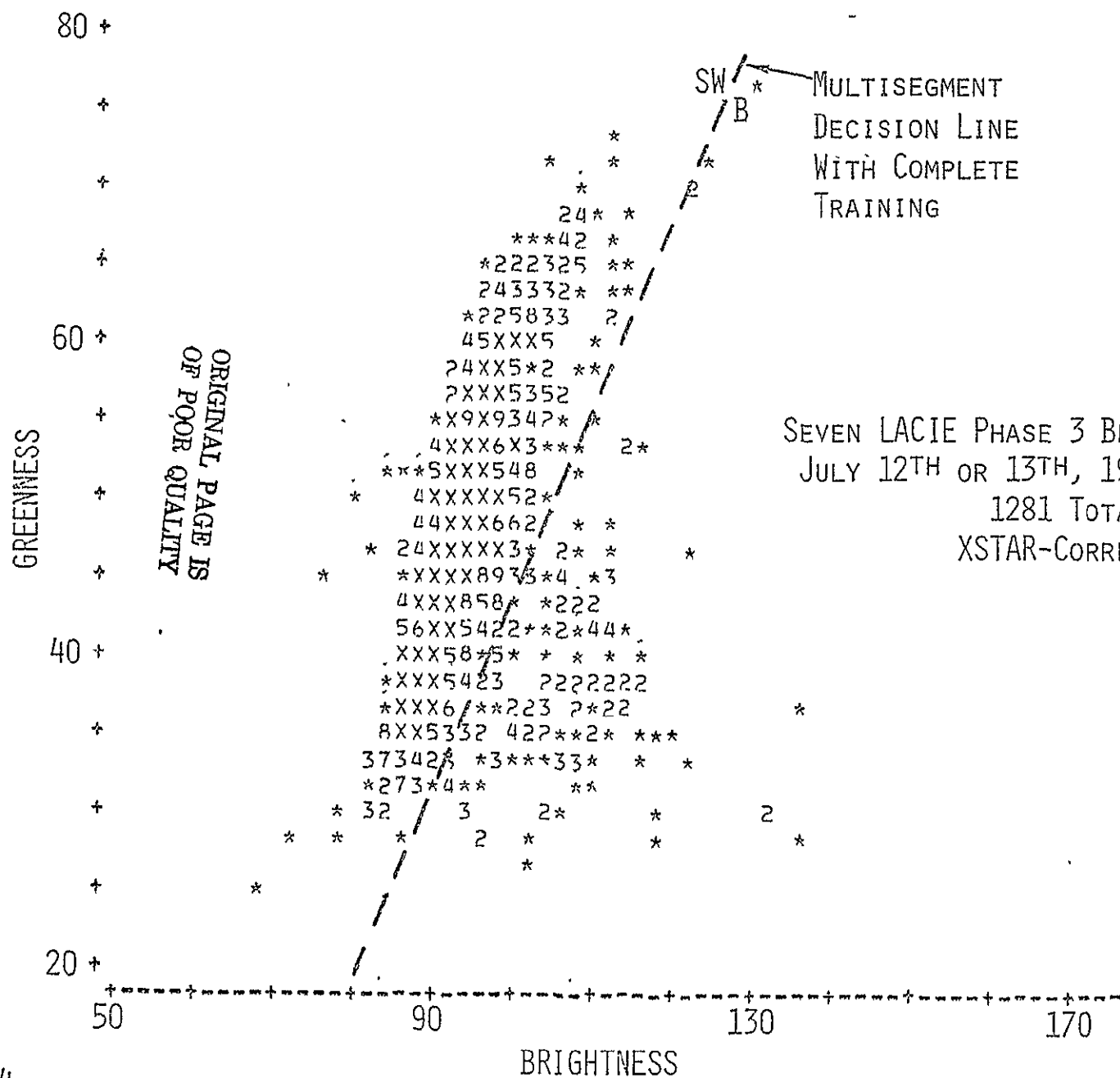
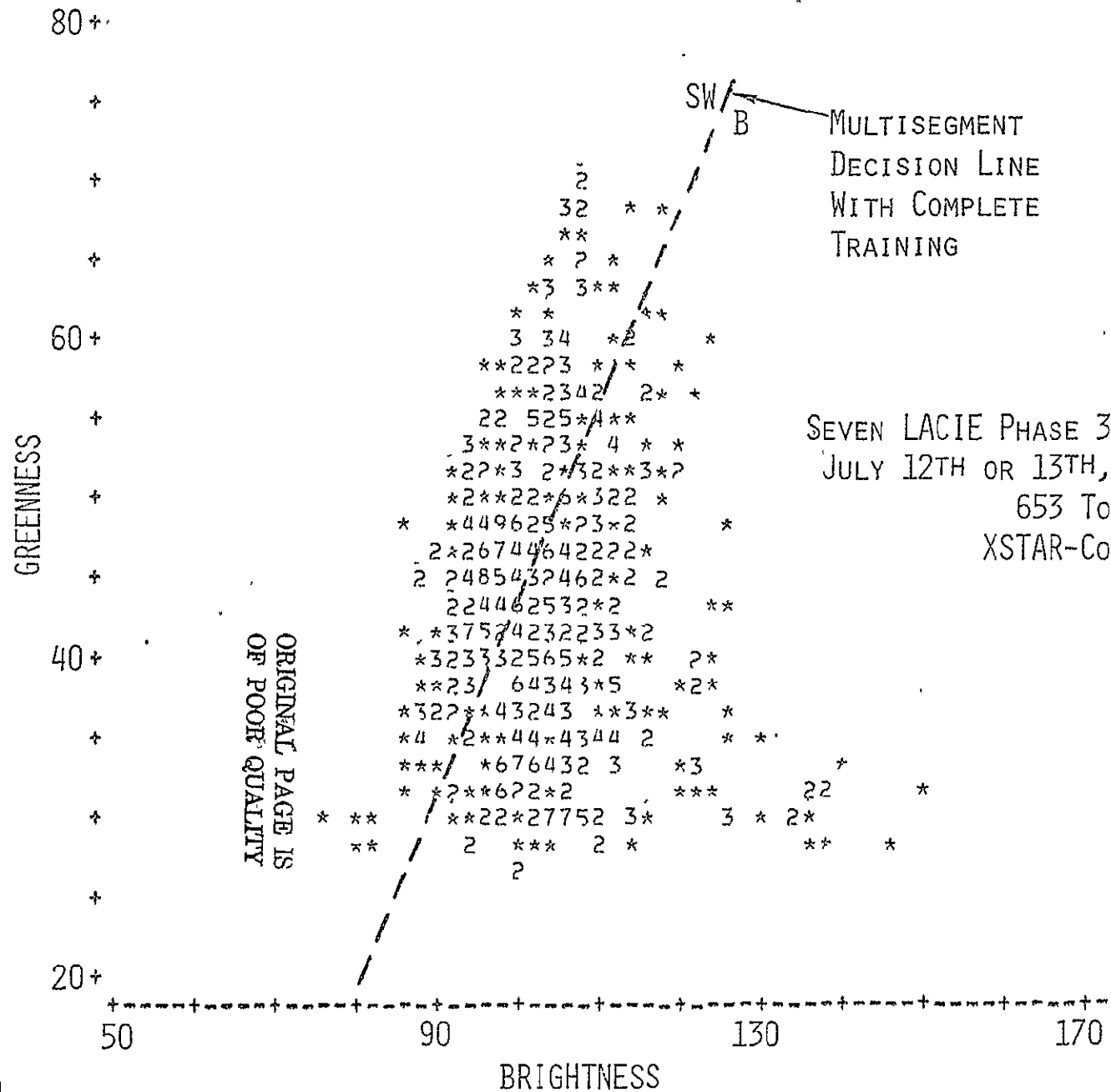


FIGURE 4.3
SCATTER OF BARLEY DATA

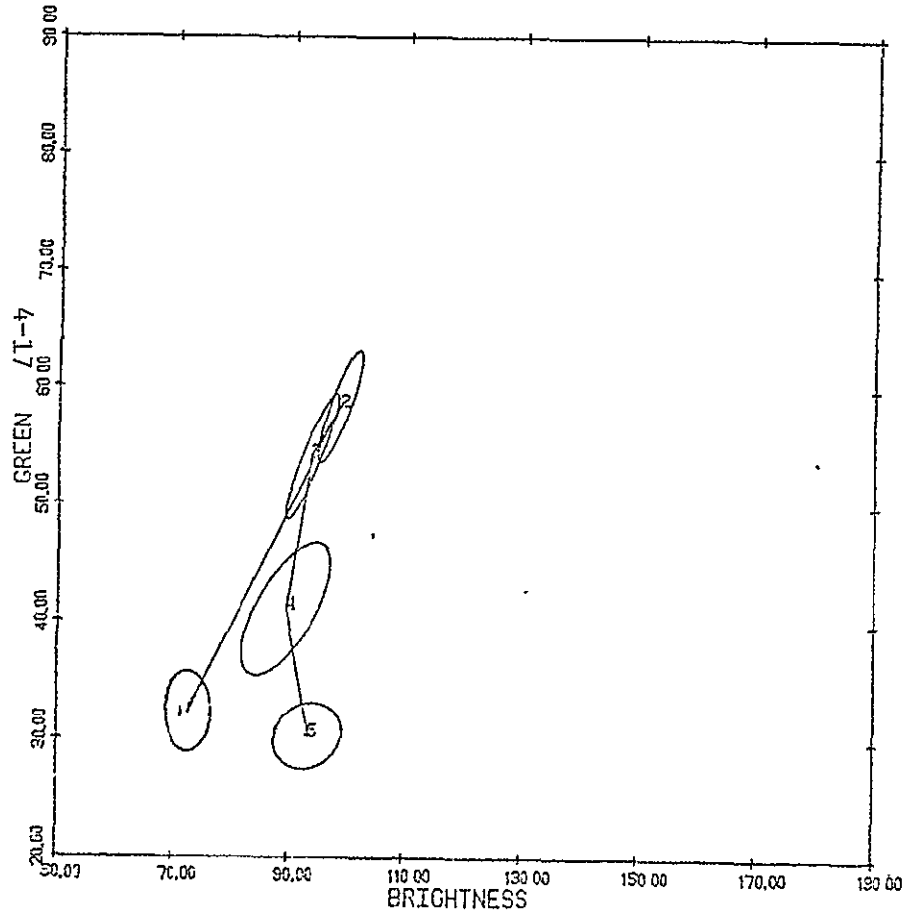


4-16

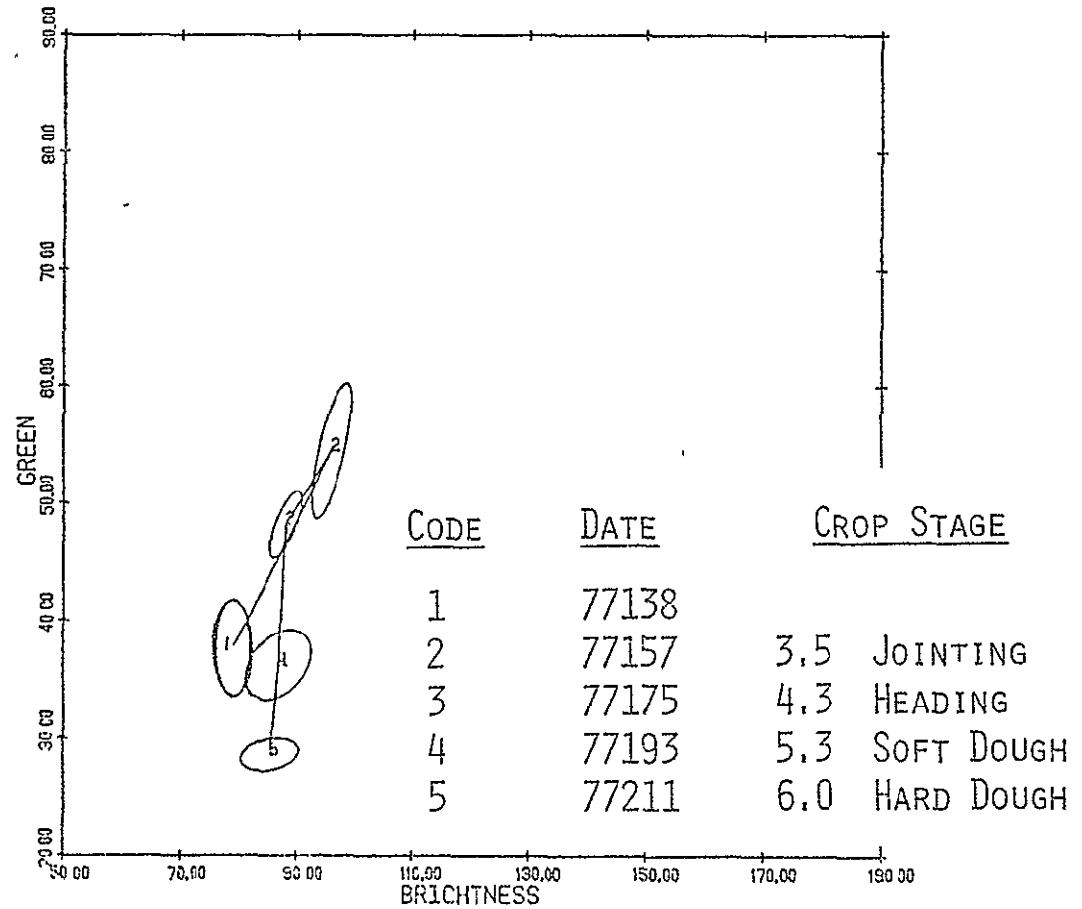
FIGURE 4.4

TEMPORAL TRACKS OF SPRING WHEAT CLUSTERS (SEGMENT 1663)

WHEAT CLUSTER TRAJECTORIES 5 DATES



WHEAT CLUSTER TRAJECTORIES 5 DATES

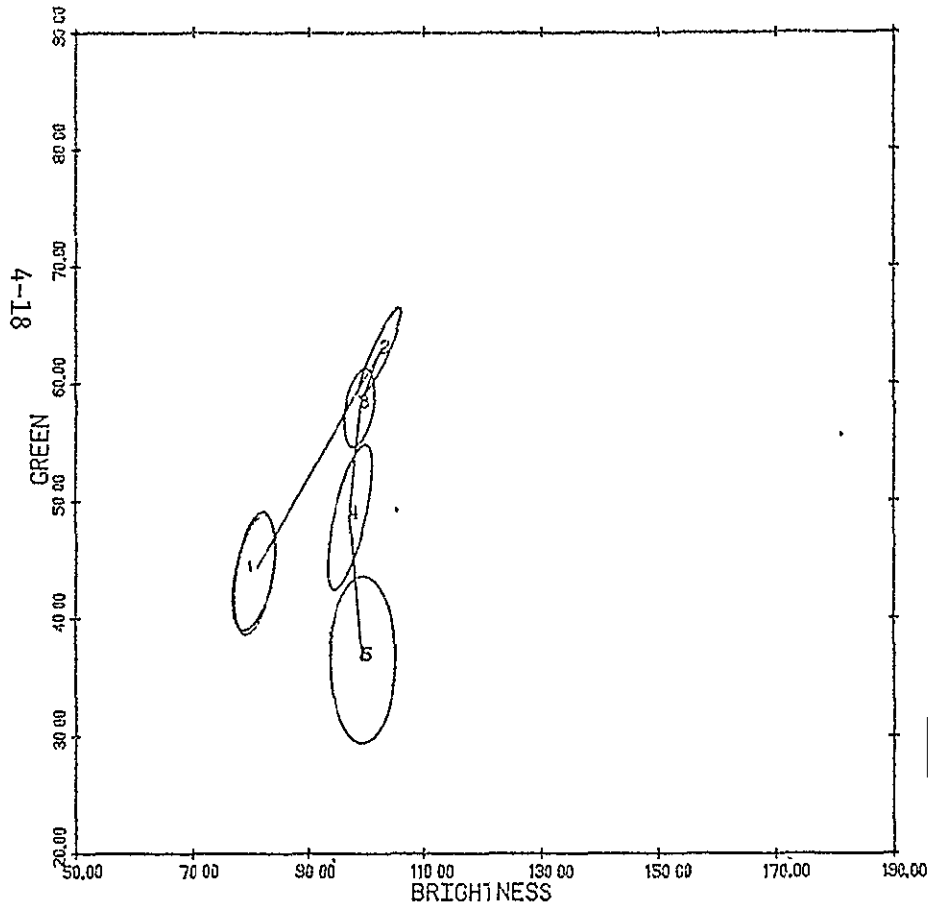


TASK 4

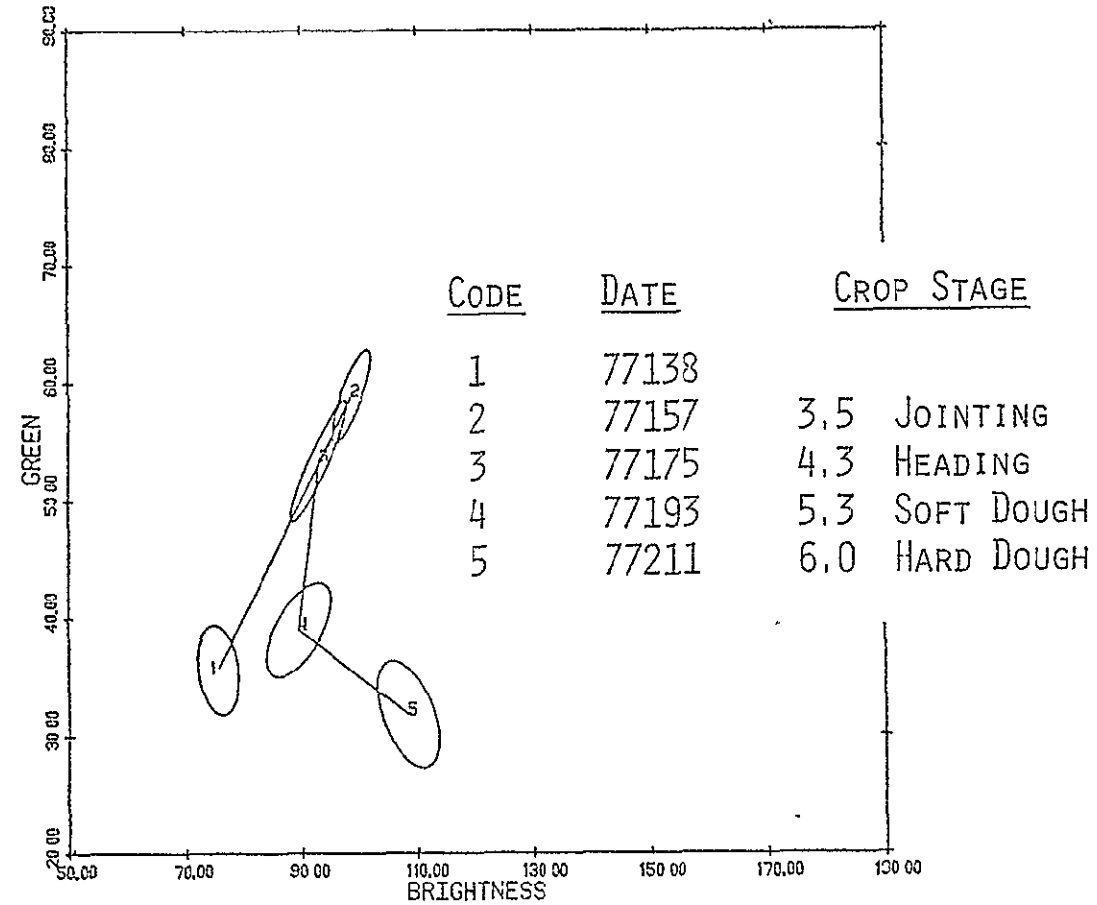
FIGURE 4.4

TEMPORAL TRACKS OF SPRING WHEAT CLUSTERS (SEGMENT 1663) (CONT.)

WHEAT CLUSTER TRAJECTORIES 5 DATES



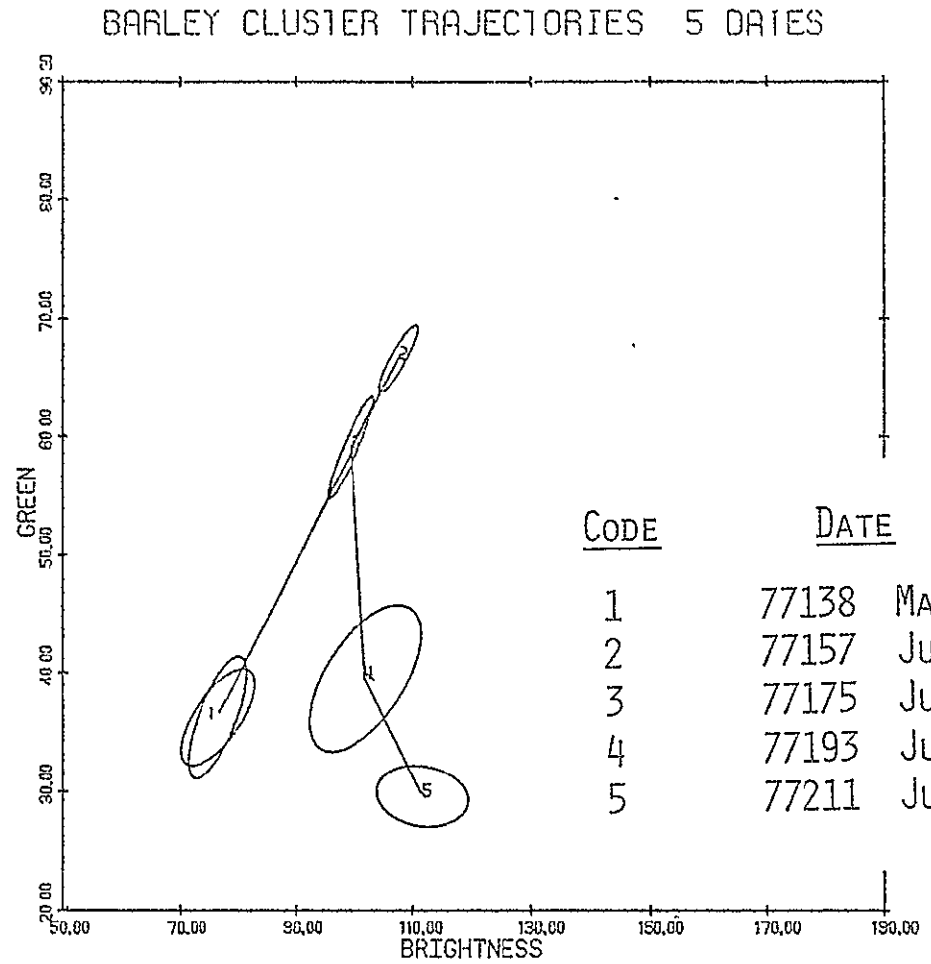
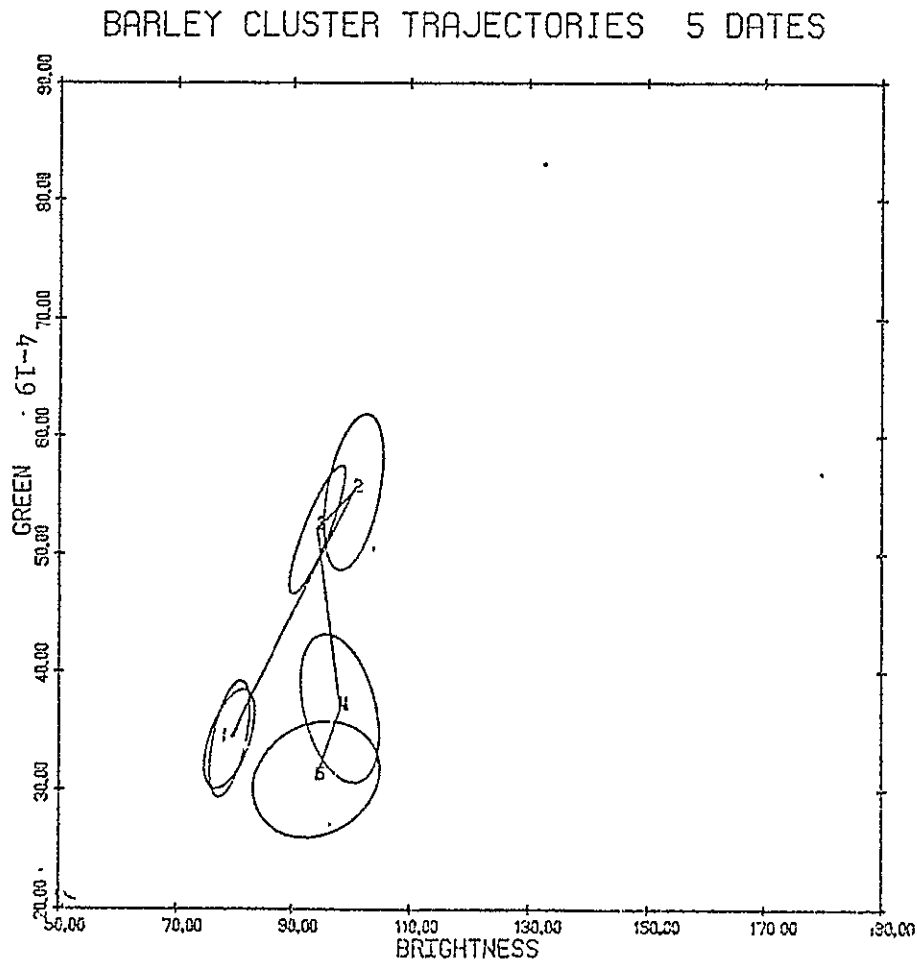
WHEAT CLUSTER TRAJECTORIES 5 DATES



TASK 4

FIGURE 4.5

TEMPORAL TRACKS OF BARLEY CLUSTERS (SEGMENT 1663)



TASK 4

TASK 6

FORECASTING PRODUCTION OF WHEAT FROM SATELLITE DATA

(R.F. Nalepka)

6.1 OBJECTIVE

The objective of this task is to investigate the utility of Landsat, meteorological, and ancillary data for forecasting winter wheat yield.

6.2 APPROACH

The approach to fulfilling the objective of this task is to examine two separate topics related to the general issue of wheat production forecasting. The principal topic to be investigated is the construction and implementation of improved and combination yield models. Another topic to be investigated is the generality and large area applicability of various yield forecasting approaches. In order to satisfactorily investigate the two topics we are also developing a larger data base with which to test various approaches.

6.3 TECHNICAL DISCUSSION

Our current technical activities are described in Figure 6.1. Details of the generation of new data are shown in Figures 6.2 and 6.3. The preliminary data analysis that has been done is shown in Figure 6.4.

Figure 6.5 shows a comparison of Landsat, CCEA, and early KCLRS forecasts of yield for counties in the southwest Crop Reporting District (CRD). KCLRS estimates of 1976 yield made in March of 1977 are used as the standard of comparison.

Figure 6.6 shows a stepwise regression analysis of the relative utility of various factors for estimating yield in 13 counties of southwest Kansas.

The results involving Landsat estimates of green measure and yield reported here should be considered as preliminary in nature. The optimal procedure to get an accurate estimate of Landsat green measure and yield is not yet clear, and there are indications of some error (although compensating) in the methods we used to generate this data. A more controlled method would be to use data from known wheat fields, on blind sites. However, it is not yet clear that such a procedure will produce a representative sample of wheat conditions.

It has been proposed that one factor limits crop yield in any particular situation, and that this concept could provide the basis for a yield model that is superior in some respects to multiple linear regression. A law-of-minimum (LOM) model can be constructed by determining relationships between separate factors and yield under conditions identified as limiting to each factor. The model can be implemented by calculating predicted yields from all of the above relationships, and choosing the minimum value for each observation as the correct prediction of yield.

We have constructed LOM models on data from 13 counties in southwest Kansas using two different procedures, and compared these models with a model generated using multiple linear regression. The results are shown in Figure 6.7.

These results suggest that an LOM model may be better or worse than multiple linear regression, depending on exactly how the model was constructed. Our preliminary conclusion is that construction of an optimum LOM model probably can't be consistently achieved using a single unsupervised algorithm. Rather, considerable human interaction, judgement, and perhaps trial and error, may be required.

Figure 6.1

CURRENT ACTIVITIES

- GENERATE DATA SET OF LANDSAT, METEOROLOGICAL AND CULTURAL VARIABLES FOR SOUTHWEST KANSAS, SIMILAR TO DATA SET PREVIOUSLY ESTABLISHED FOR CENTRAL CRD.
- PRELIMINARY ANALYSIS OF NEW DATA (INCLUDING CATE & LIEBIG)

TASK 6

Figure 6.2
GENERATION OF LANDSAT DATA

- MAKE ESTIMATES OF LANDSAT GREEN MEASURE ON,
COUNTY-BY-COUNTY BASIS
- MAKE LANDSAT ESTIMATES OF YIELD ON COUNTY-BY-COUNTY
BASIS USING FINNEY ITS DATA FOR TRAINING

TASK 6

Figure 6.3

GENERATION OF ANCILLARY DATA SET

- ORGANIZE MONTHLY METEOROLOGICAL DATA
- COMPUTE CCEA YIELD ESTIMATES
- CALCULATE KCLRS YIELD ESTIMATES
- DETERMINE CULTURAL DATA (% SUMMER FALLOW, % IRRIGATED)

TASK 6



Figure 6.4

PRELIMINARY ANALYSIS OF DATA

- COMPARE LANDSAT AND CCEA ESTIMATES OF YIELD WITH KCLRS YIELD
- GENERATE LAW-OF-MINIMUM MODELS FOR COUNTIES OF SOUTHWEST KANSAS
USING LANDSAT GREEN MEASURE, AUG - FEB PRECIPITATION, MAY PRECIPITATION,
PERCENT SUMMER FALLOW, PERCENT IRRIGATED
- GENERATE LINEAR REGRESSION MODEL USING ABOVE DATA
- COMPARE LOM AND REGRESSION MODELS

TASK 6

Figure 6.5

COMPARISON OF SEPARATE LANDSAT AND CCEA YIELD ESTIMATES FOR
COUNTIES IN THE SOUTHWEST CRD
(COMPARED TO MARCH 1977 KCLRS ESTIMATES).

<u>ESTIMATE</u>	<u>MEAN</u>	<u>MEAN DIFF</u> FROM KCLRS
CCEA (1 JULY 1976)	30.0	+3.0
LANDSAT (18 APRIL 1976)	28.5	+1.5
KCLRS (APRIL 1976)	17.0	- 10.0

Figure 6.6

RELATIVE UTILITY OF VARIOUS FACTORS FOR ESTIMATING YIELD
(13 COUNTIES OF SW KANSAS ASSESSED BY STEPWISE REGRESSION)

STEP	VARIABLE	R ²	S.E.
1	%SUM FAL	.30	2.34
2	AVE GRN	.62	1.82
3	% IRRIG	.74	1.59
4	MAY PRE	.76	1.60
5	AUG-FEB PRE	.77	1.67

TASK 6



Figure 6.7
COMBINATION YIELD MODELS

(BASED ON 13 COUNTIES IN SW KANSAS)

<u>MODEL</u>	<u>R²</u>	<u>S.E.</u>
MULTIPLE LINEAR REGRESSION	0.76	1.6
LAW-OF-MINIMUM (FIT R)	0.65	1.7
LAW-OF-MINIMUM (FIT M)	0.86	1.0

TASK 7

PREPILOT STUDY OF MULTICROP SPECTRAL SEPARABILITY

(W.A. Malila and E.P. Crist)*

7.1 INTRODUCTION

Through a re-organization of efforts, making them better focussed on LACIE transition objectives, this task was formed and is being reported here for the first time. The key issue being addressed is the extension of large area crop inventory technology to include important crops in addition to wheat.

7.2 OBJECTIVE

The objective is to conduct preflight signature studies using currently available data, to provide insights and identify potential problem areas for investigation in future multicrop pilot testing and R&D activities by JSC and SR&T contractors.

7.3 APPROACH

The approach being taken toward meeting the objective is outlined in Table 7.1. Note the emphasis on use of currently available data.

7.4 SUMMARY OF PROGRESS DURING THE QUARTER

Table 7.2 summarizes our progress during the reporting period.

7.5 DETAILS OF PROGRESS

7.5.1 ANALYSIS OF CITARS DATA

Table 7.3 describes the CITARS data set which was the primary focus of our first quarter's activity. The locations of the sites are shown in Figure 7.1. As indicated in Table 7.4, both field-center pixels and Landsat mean vectors formed from them are being analyzed.

As our first step in investigating corn and soybeans signature characteristics, we conducted a series of unitemporal and multitemporal

* W. Malila is task leader; contributions to the reported work were also made by J.M. Gleason.

clustering operations on four-date data from the Fayette, Illinois, segment. The results are presented next in the form of spectral signature plots of Landsat Band 6 vs. Band 5. The dispersion of pixel values within each cluster is indicated by a "one-sigma" ellipse about the mean value.

The first series of plots is of multitemporal (four-date) clusters for corn and soybeans. All clusters with more than five pixels are illustrated. Figures 7.2 and 7.3 show the way corn pixels progress from a June 10th dispersion along a line, indicating much visible bare soil (of varied reflectance), to compact distributions on July 17th and August 21st. In contrast, the soybeans clusters shown in Figures 7.4 and 7.5 are widely dispersed on all dates, with most compactness on the last date, August 21. A contributing factor is the fact that some soybeans were planted several weeks late in fields from which winter wheat had been harvested.

The patterns for these crops are shown in more detail on a unitemporal basis in Figures 7.6 through 7.9; all clusters with ten or more pixels are shown. Clustering for the latter two figures included pixels from other crops as well as corn and soybeans.

Finally, three-date multitemporal clusters, generated using all labeled pixels, are presented in Figures 7.10 and 7.11.

From these initial results, we do not attempt to draw conclusions regarding the best date for classifying corn and soybeans or the accuracy levels to be expected. However, we do note that substantial spectral differences do exist, that the distribution of corn on the last two dates is compact, and that multitemporal clusters were more pure than unitemporal clusters.

In addition to the preceding Landsat data analysis, we have digitized the periodic (18-day) observation data from the Fayette segment, to enable subsequent comparisons of spectral data and agronomic characteristics of the fields.



7.5.2 ANALYSIS OF LACIE PHASE 3 BLIND SITE DATA

To increase the data base for corn and soybeans, we have prepared data from eleven LACIE Phase 3 Blind Sites which contain some corn or soybean fields. The data preparation steps and amounts of data available are summarized in Table 7.5.

7.6 PLANS

Plans for the next quarter are presented in Table 7.6.

TABLE 7.1

APPROACH

- ANALYZE AVAILABLE LANDSAT DATA ON CORN AND SOYBEANS
 - 1973 CITARS DATA (ILLINOIS AND INDIANA)
RE-EXAMINE IN LIGHT OF UNDERSTANDING AND NEW TECHNIQUES
GAINED THROUGH LACIE:
 - STRUCTURE OF LANDSAT DATA (TASSELLED CAP TRANSFORMATION)
 - SPECTRAL/SPATIAL AND MULTITEMPORAL CLUSTERING
 - XSTAR PREPROCESSING (CORRECTIONS FOR SUN ANGLE AND HAZE)
 - QUANTITATIVE USE OF PERIODIC GROUND OBSERVATIONS
 - 1977 LACIE BLIND SITE DATA (SELECTED SITES ON FRINGES OF A
U.S. CORN BELT)
- ANALYZE AVAILABLE DATA FROM OTHER SOURCES
 - FIELD MEASUREMENTS OF SPECTRAL REFLECTANCE
 - SIMULATION MODEL DATA

TABLE 7.2

SUMMARY OF PROGRESS

- HAVE BEGUN RE-EXAMINATION OF CITARS DATA IN LIGHT OF RECENT TECHNOLOGY
 - LANDSAT SIGNATURES AS FUNCTION OF TIME
 - MULTITMPORAL CLUSTERING OF FAYETTE SEGMENT
 - QUANTITATIVE ANALYSIS OF PERIODIC GROUND OBSERVATIONS
- HAVE COMPLETED PREPARATION OF ELEVEN PHASE 3 BLIND SITES CONTAINING SOME CORN AND SOYBEANS FOR MULTICROP ANALYSIS
- HAVE REQUESTED EARLIER FIELD MEASUREMENTS OF CORN AND SOYBEAN SPECTRAL REFLECTANCE DATA FROM LARS

TABLE 7.3
CITARS DATA SET

• LANDSAT DATA

- ACQUIRED THROUGHOUT SUMMER OF 1973
- SIX 5x20-MILE SEGMENTS IN INDIANA AND ILLINOIS
- CLOUD COVER SUBSTANTIALLY REDUCED THE NUMBER OF USEABLE ACQUISITIONS, FAYETTE, ILL., SEGMENT HAD THE MOST, USEABLE ACQUISITIONS (EARLY AND LATE JUNE, MID JULY, AND LATE AUGUST)
- COLOR COMPOSITE IMAGES RECENTLY PRODUCED ON THE LACIE PRODUCTION FILM CONVERTER (PFC)
- LACIE-SIZED (5x6-MI) SEGMENTS RECENTLY EXTRACTED FROM THE 5x20-MILE SEGMENTS

• FIELD IDENTIFICATION INFORMATION (~25% OF AREA)

- 20 QUARTER SECTIONS BY PERIODIC GROUND OBSERVATIONS.
- 20 FULL SECTIONS BY INTERPRETATION OF MULTIDATE AERIAL PHOTOGRAPHS

• AGRONOMIC DATA

- PERIODIC (18-DAY) GROUND OBSERVATIONS OF PLANT HEIGHT, GROUND COVER, CROP MATURITY STAGE, AND CONDITION
- FIELD SIZE, ROW WIDTH, AND ROW DIRECTION

TABLE 7.4

LANDSAT DATA ANALYZED

- PIXELS WITHIN IDENTIFIED FIELDS
 - BOUNDARY OR EDGE PIXELS WERE EXCLUDED TO ELIMINATE THIS SOURCE OF CONFUSION FROM INITIAL STUDIES
 - CROP LABEL AND FIELD NUMBER WERE APPENDED TO EACH PIXEL IN FAYETTE SEGMENT
 - MULTITEMPORAL DATA SET ASSEMBLED FOR FAYETTE SEGMENT
- LANDSAT MEAN VECTOR FROM EACH OF THESE FIELD CENTERS
 - MEAN VECTORS USED IN CITARS WERE ANALYZED
 - ALL SEGMENTS AND ALL USEABLE DATES

7-7

6-2

TABLE 7.5

LACIE PHASE 3 BLIND SITE DATA

- ELEVEN SEGMENTS PREPARED FOR ANALYSIS
(MOST INCLUDED IN JSC'S LIST OF CANDIDATE SEGMENTS FOR MULTICROP ANALYSES.)
 - SCREENED (VISUALLY AND BY COMPUTER ALGORITHM)
 - XSTAR PREPROCESSED (SUN ANGLE AND HAZE CORRECTIONS)
 - EXTRACTED 897 POINTS (5x5 GRID; INCLUDES 209 DOTS AS SUBSET)
 - MERGED WITH GROUND TRUTH DATA FROM JSC TAPES
(LABELED AS EITHER A SINGLE SPECIFIC COVER CODE OR A MIXTURE PIXEL.)

7-8

- SEGMENTS, WITH CORN AND SOYBEANS PERCENTAGES (FROM 897 POINTS)

<u>STATE</u>	<u>SEGMENT</u>	<u>% CORN</u>	<u>% SOYBEANS</u>	<u>STATE</u>	<u>SEGMENT</u>	<u>% CORN</u>	<u>% SOYBEANS</u>
ND	1619	2.3 (21)*	0	SD	1669	1.2 (11)	0
	1640	3.2 (29)	0		1686.	10.9 (98)	0
	1663	0.4 (4)	5.4 (48)		1699	2.6 (23)	0
	1903.	2.6 (23)	0		1802	15.5 (139)	0
	1927	5.2 (47)	0		1805.	7.6 (68)	0
MINN	1515.	4.3 (39)	3.6 (32)				

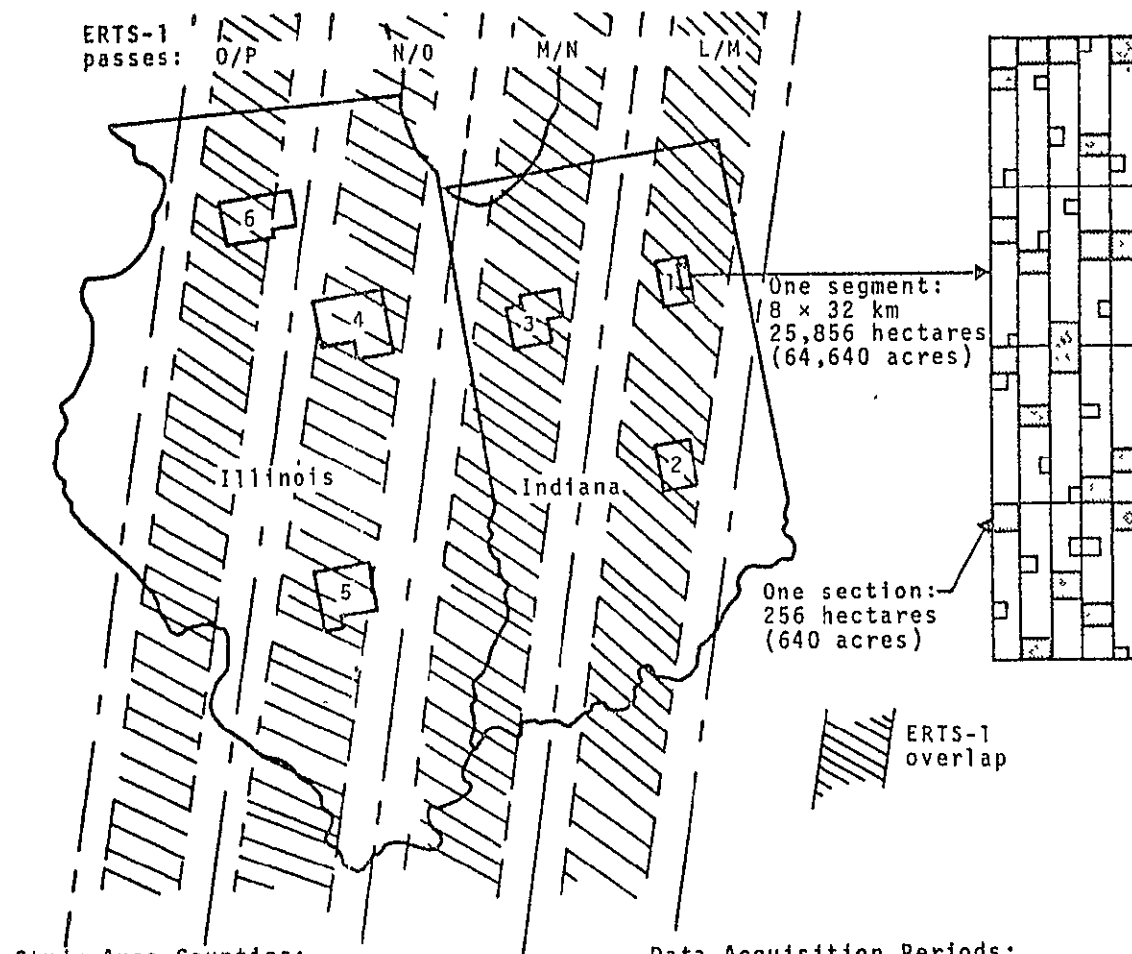
* NUMBER OF POINTS.

TABLE 7.6

PLANS

- APPLY XSTAR PREPROCESSING AND TASSELLED-CAP TRANSFORMATION TO FIELD-CENTER DATA FROM ALL SEGMENTS AND ANALYZE
(CALIBRATIONS OF THE EARLY LANDSAT-1 CITARS DATA AND THE LATER LANDSAT-1 AND LANDSAT-2 DATA HAVE BEEN COMPARED UNDER THE OTHER ERIM MULTICROP TASK.)
- PERFORM QUANTITATIVE ANALYSIS OF PERIODIC OBSERVATIONS FROM ALL SEGMENTS
 - RELATIVE CROP CALENDARS
 - RELATIVE CROP GROWTH AND CONDITION
 - RELATE TO SPECTRAL SIGNATURES
- INITIATE ANALYSIS OF SEPARABILITY IN LARGER CONTIGUOUS GROUND AREAS, INCLUDING BOUNDARY PIXELS

FIGURE 7.1
CITARS DATA SITES



Study Area Counties:

<u>Indiana</u>	<u>Illinois</u>
1. Huntington	4. Livingston
2. Shelby	5. Fayette
3. White	6. Lee

Data Acquisition Periods:

0 - 5/21-25/73	IV - 8/01-05/73
I - 6/08-12/73	V - 8/19-23/73
II - 6/26-30/73	VI - 9/06-10/73
III - 7/14-18/73	VII - 9/24-28/73

Ground Truth:

ASCS - 20 quarter sections (white) each ERTS-1 pass

Photointerpretation - 20 sections (black) each ERTS-1 pass

FIGURE 7.2

MULTITEMPORAL CLUSTER ANALYSIS - FAYETTE COUNTY

PIXELS LABELED CORN; FOUR DATES, 16 CHANNELS

PATTERN OF CORN CLUSTERS ON FIRST TWO DATES
(338 PIXELS)

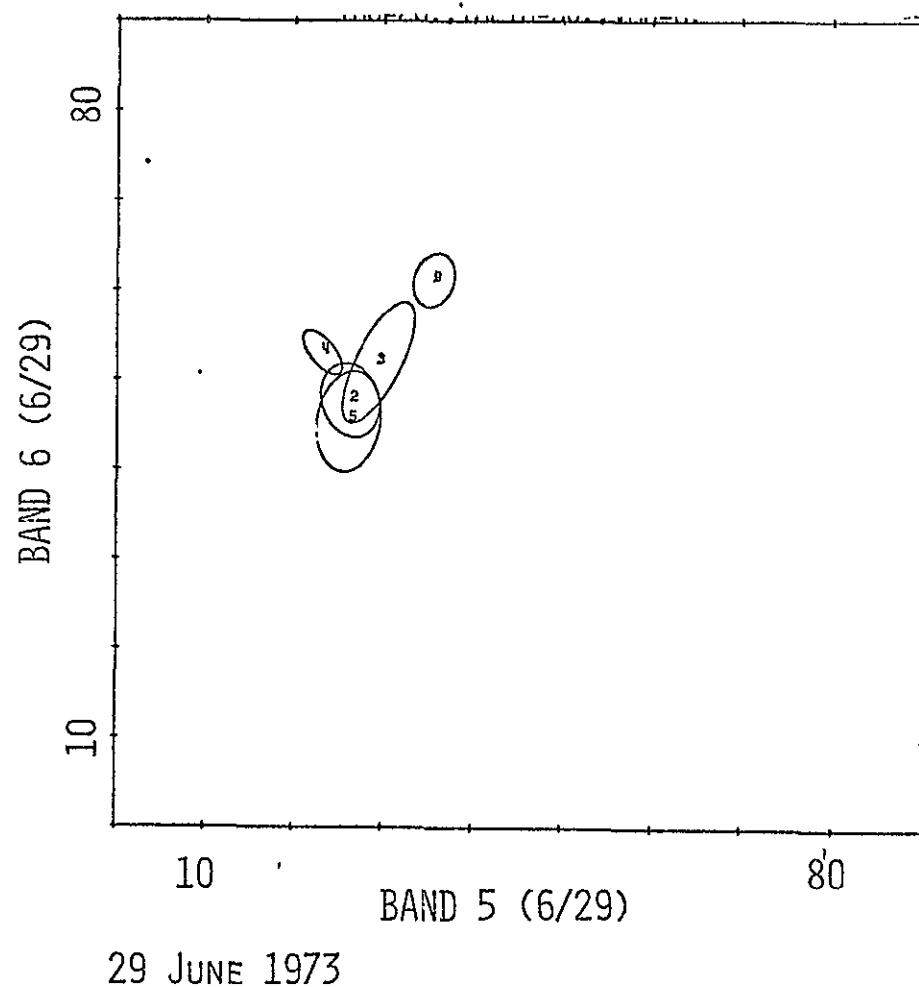
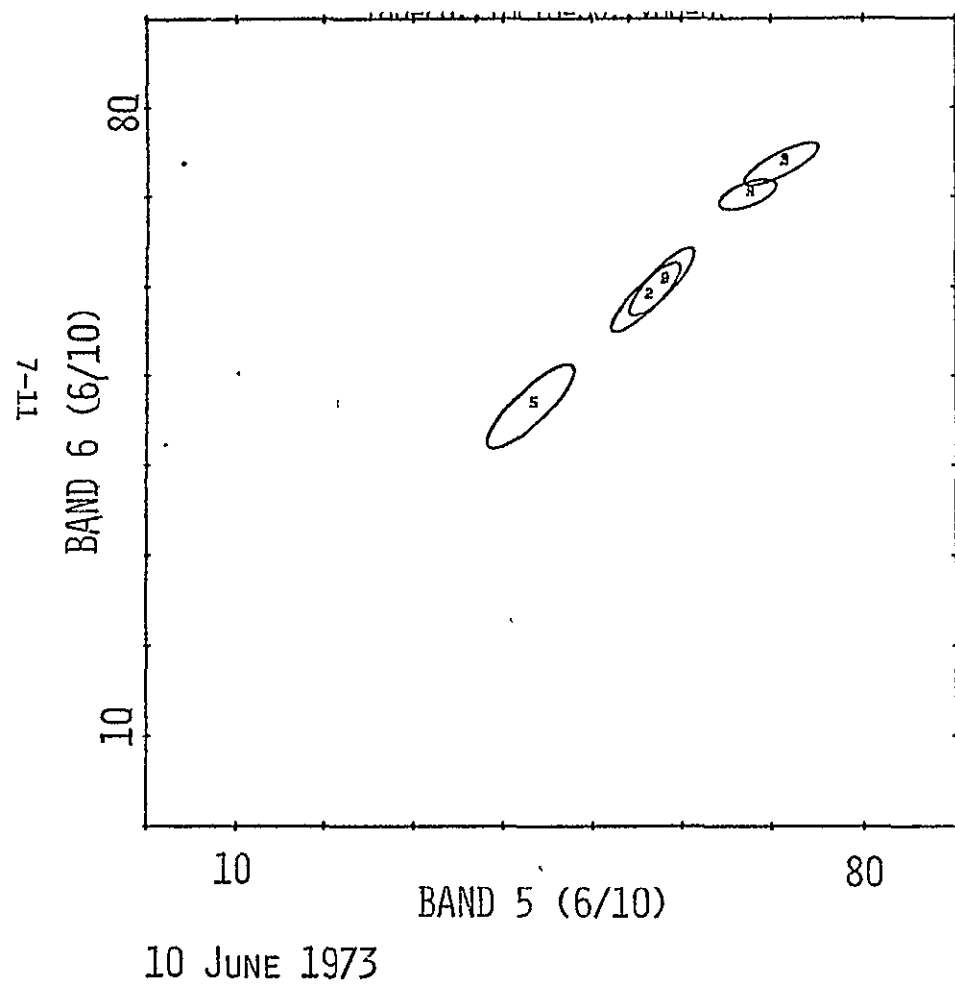


FIGURE 7.3

MULTITEMPORAL CLUSTER ANALYSIS - FAYETTE COUNTY

PIXELS LABELED CORN; FOUR DATES, 16 CHANNELS

PATTERN OF CORN CLUSTERS ON LAST TWO DATES

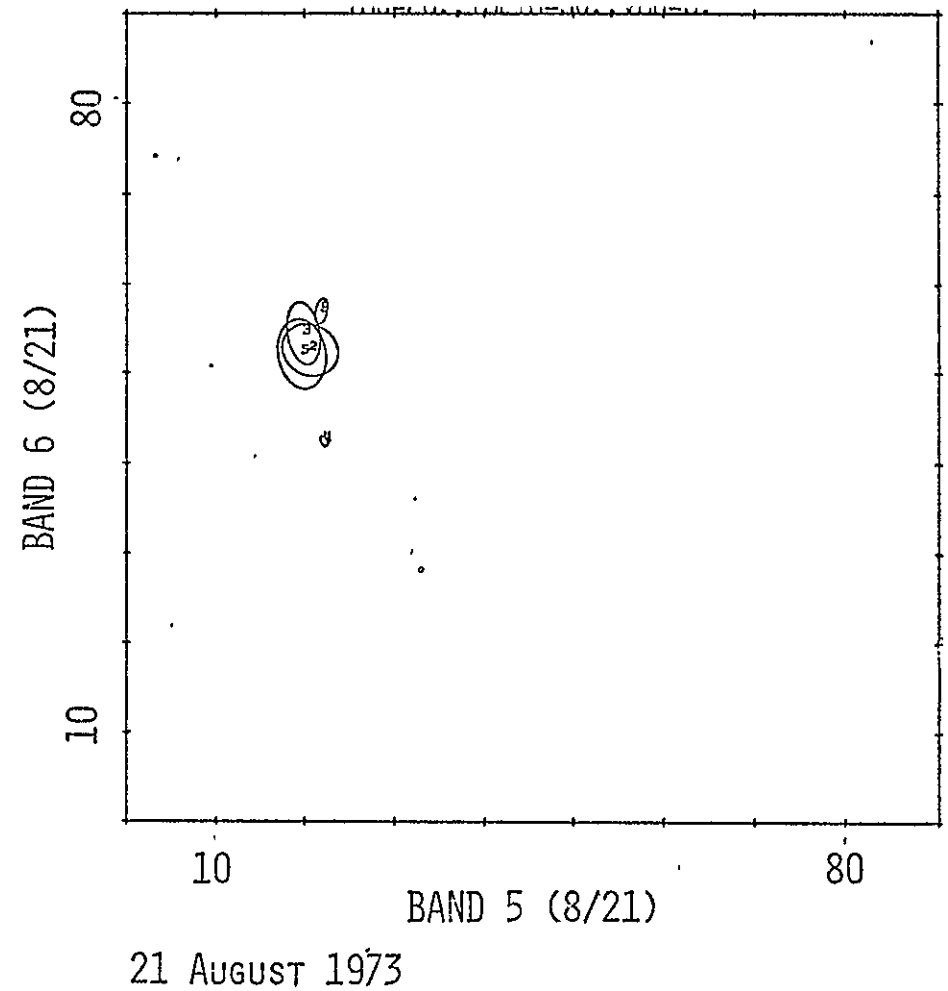
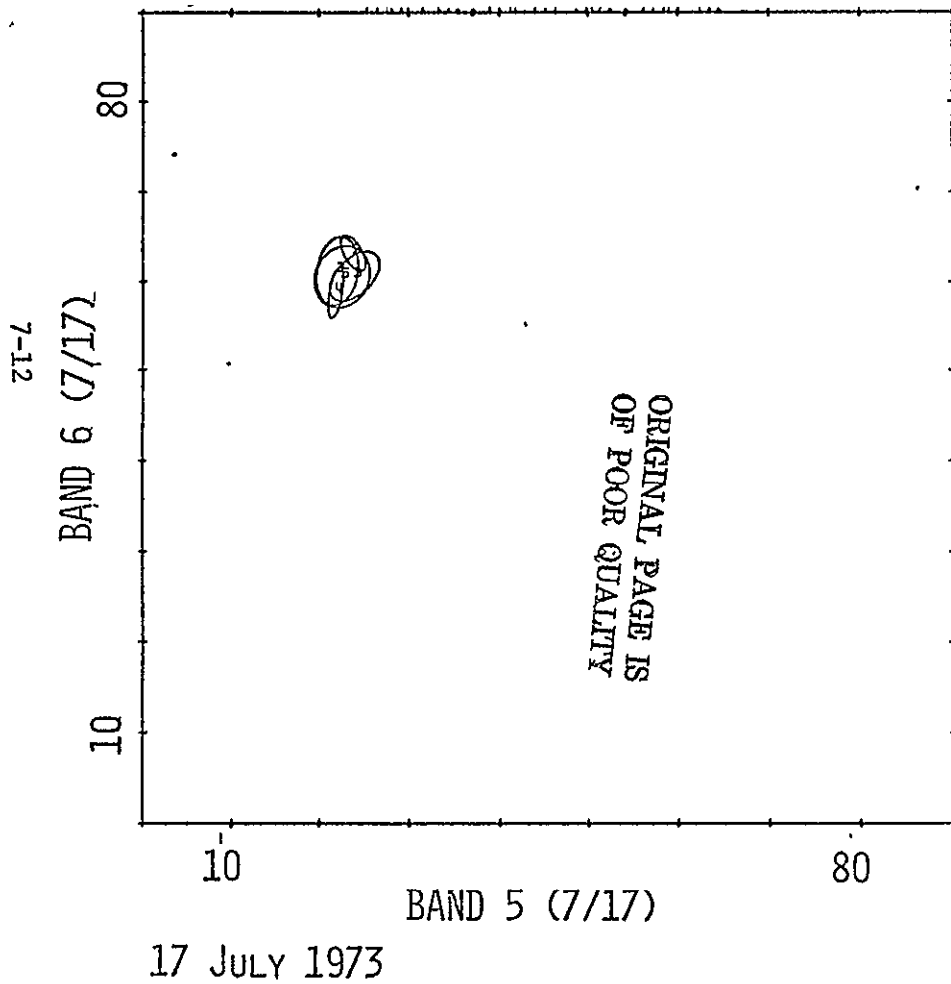


FIGURE 7.4

MULTITEMPORAL CLUSTER ANALYSIS - FAYETTE COUNTY
PIXELS LABELED SOYBEANS ON 21 AUGUST; FOUR DATES, 16 CHANNELS

PATTERN OF SOYBEANS CLUSTERS ON FIRST TWO DATES
(508 PIXELS)

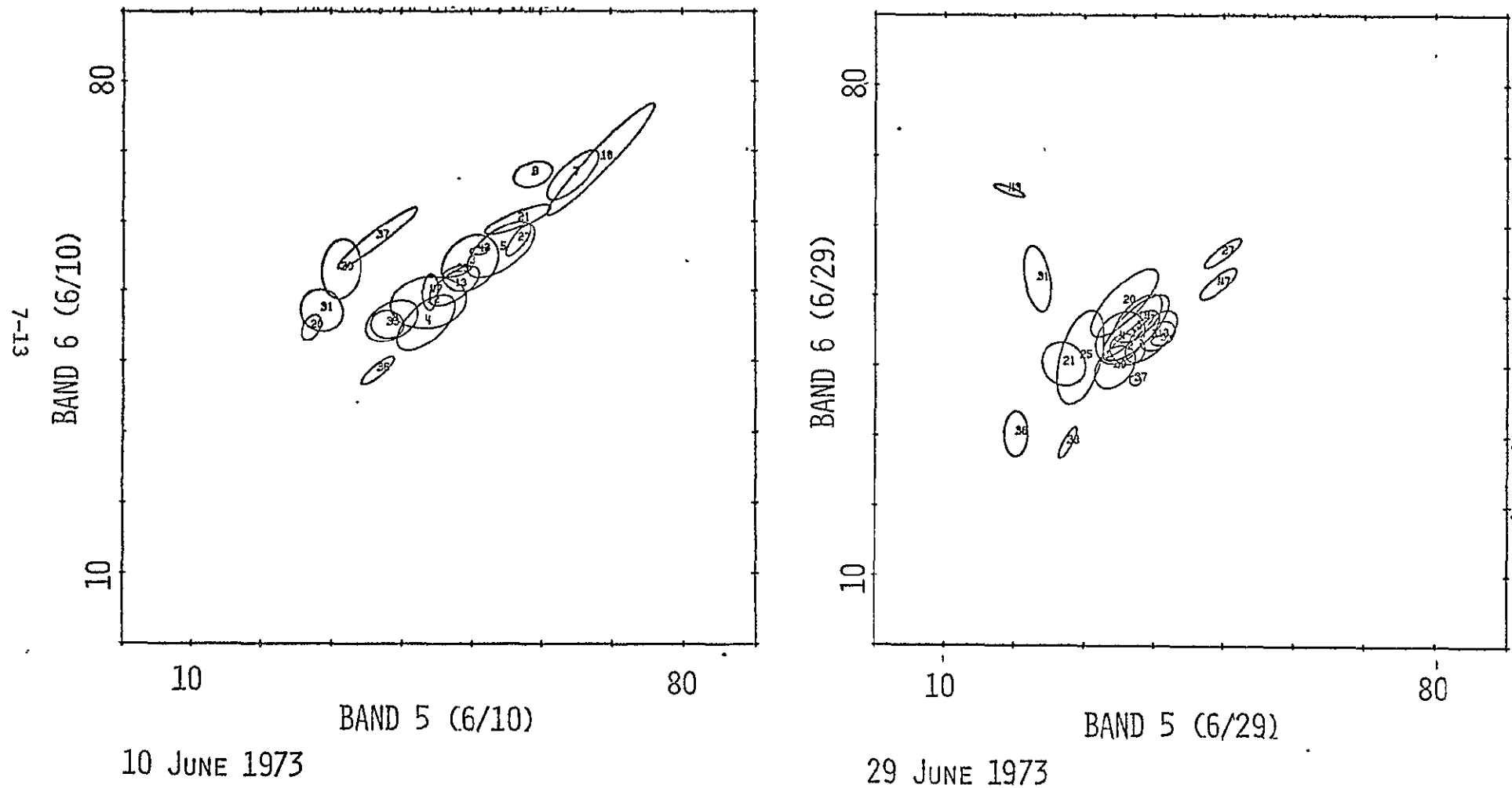
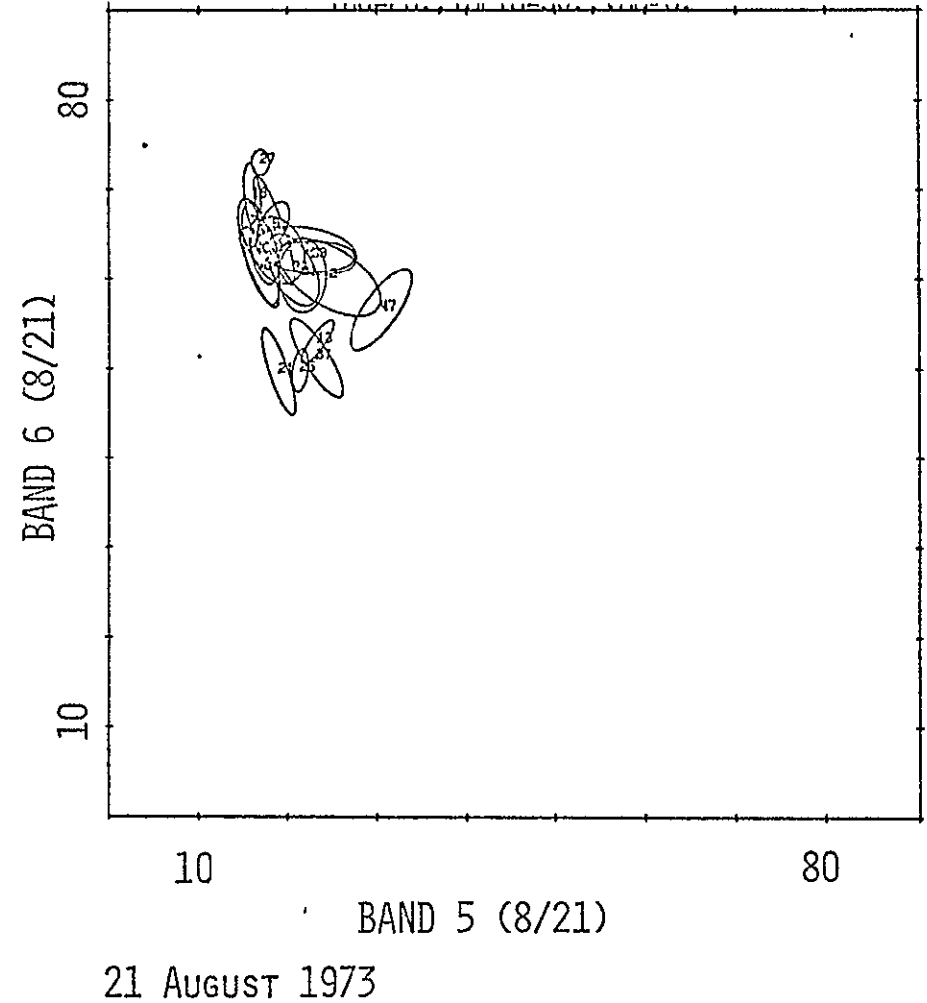
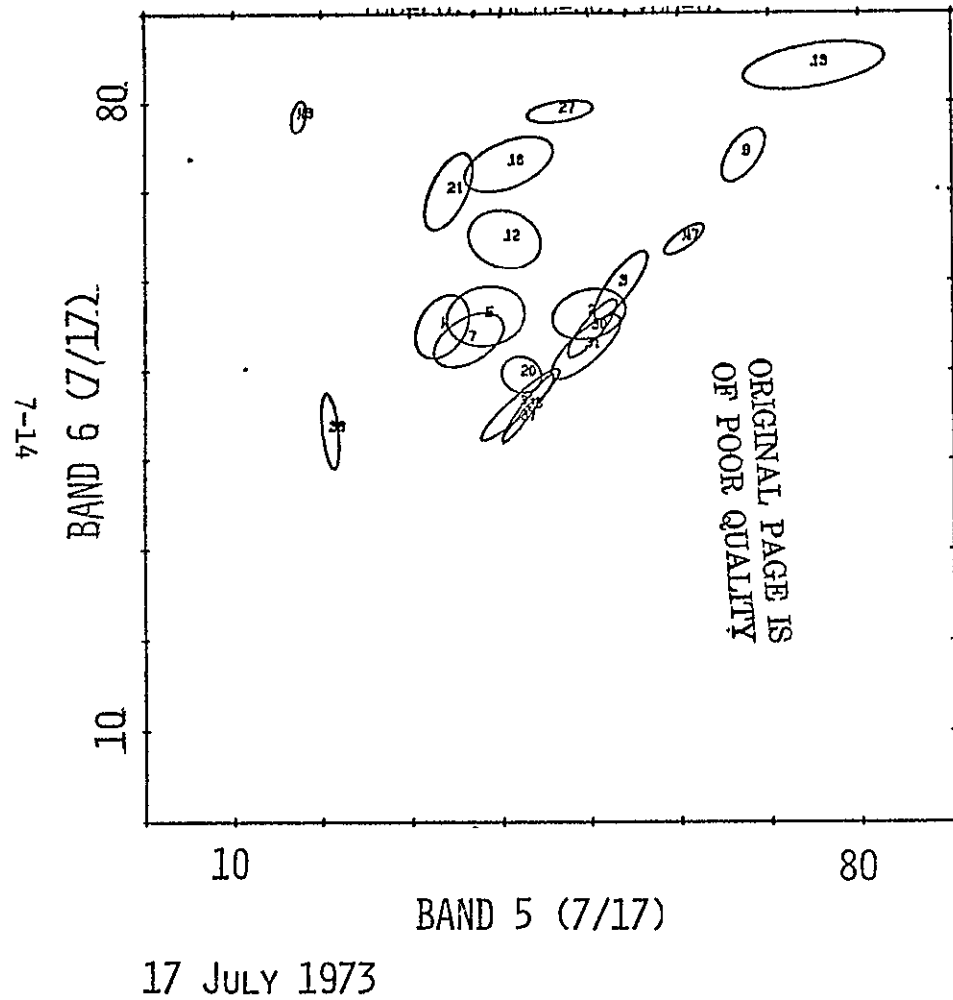


FIGURE 7.5

MULTITEMPORAL CLUSTER ANALYSIS - FAYETTE COUNTY
PIXELS LABELED SOYBEANS ON 21 AUGUST; FOUR DATES, 16 CHANNELS

PATTERN OF SOYBEANS CLUSTERS ON LAST TWO DATES



UNITEMPORAL CLUSTER ANALYSES - FAYETTE COUNTY

RESULTS USING PIXELS LABELED CORN OR SOYBEANS

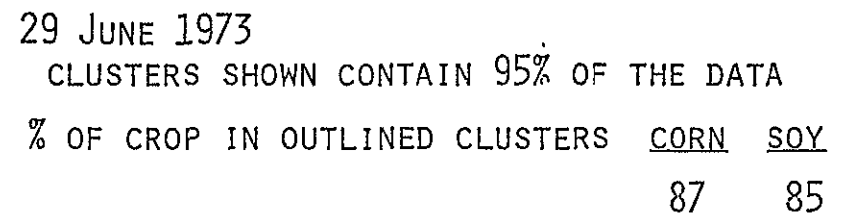
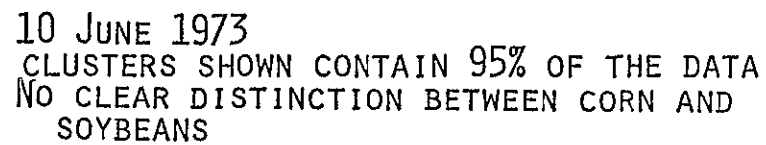
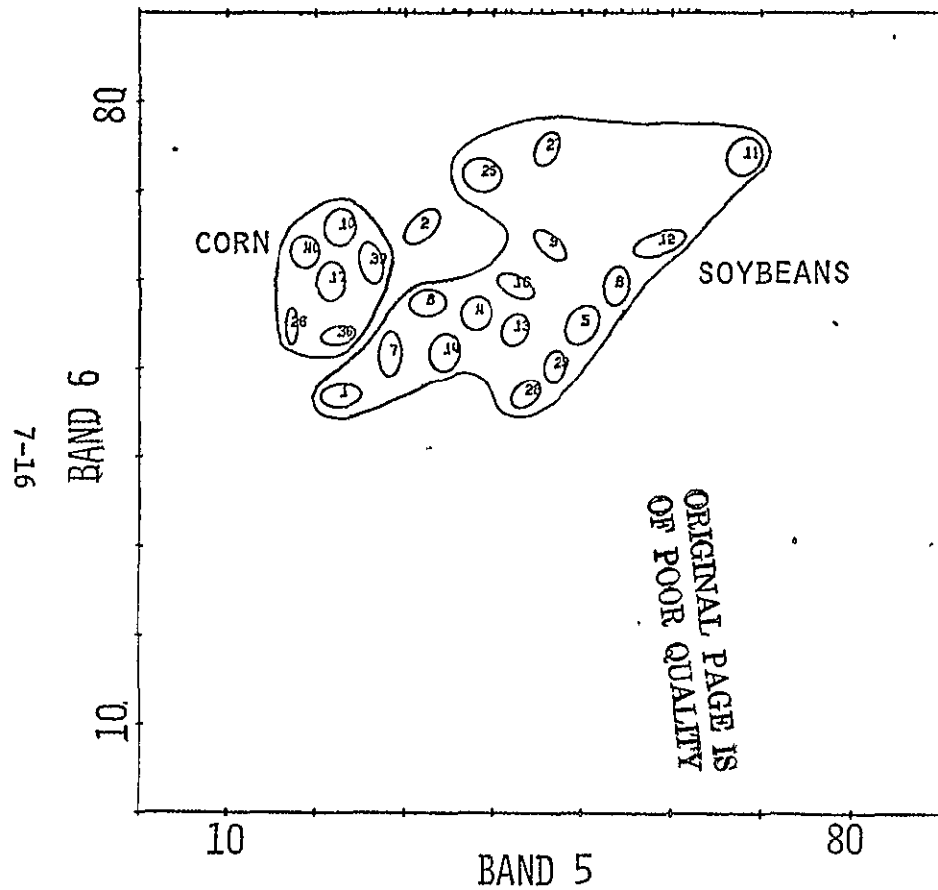


FIGURE 7.7

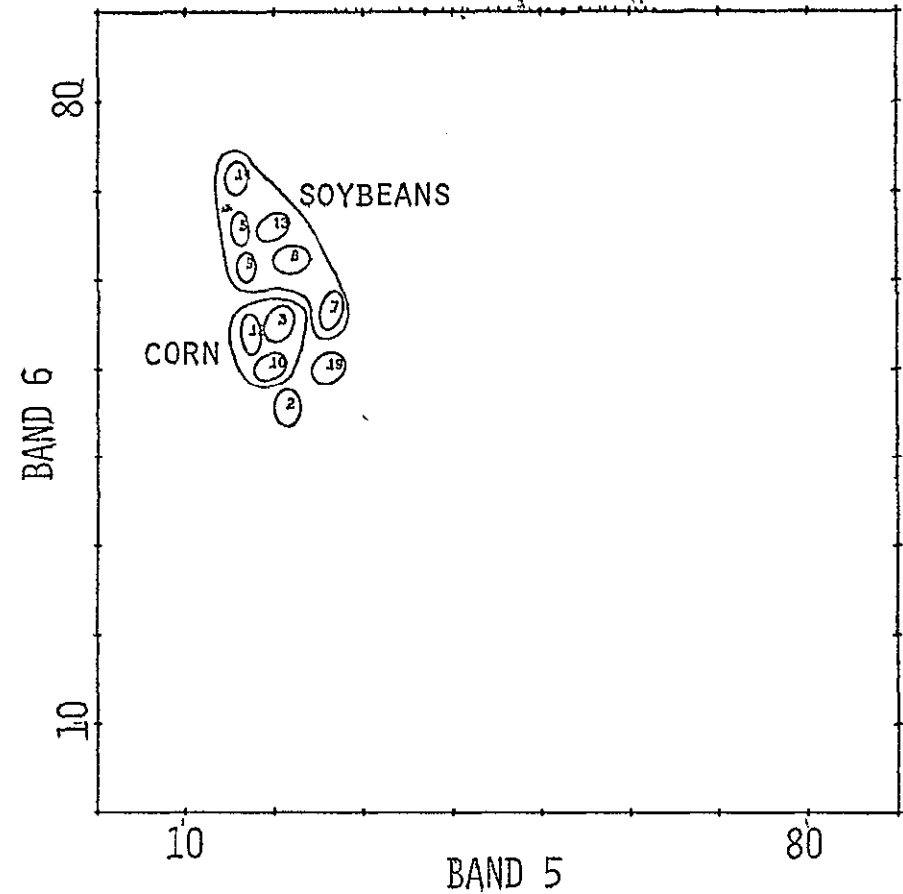
UNITEMPORAL CLUSTER ANALYSES - FAYETTE COUNTY RESULTS USING PIXELS LABELED CORN OR SOYBEANS



17 JULY 1973

CLUSTERS SHOWN CONTAIN 91% OF THE DATA

% OF CROP IN OUTLINED CLUSTERS	CORN	SOY
	91	83



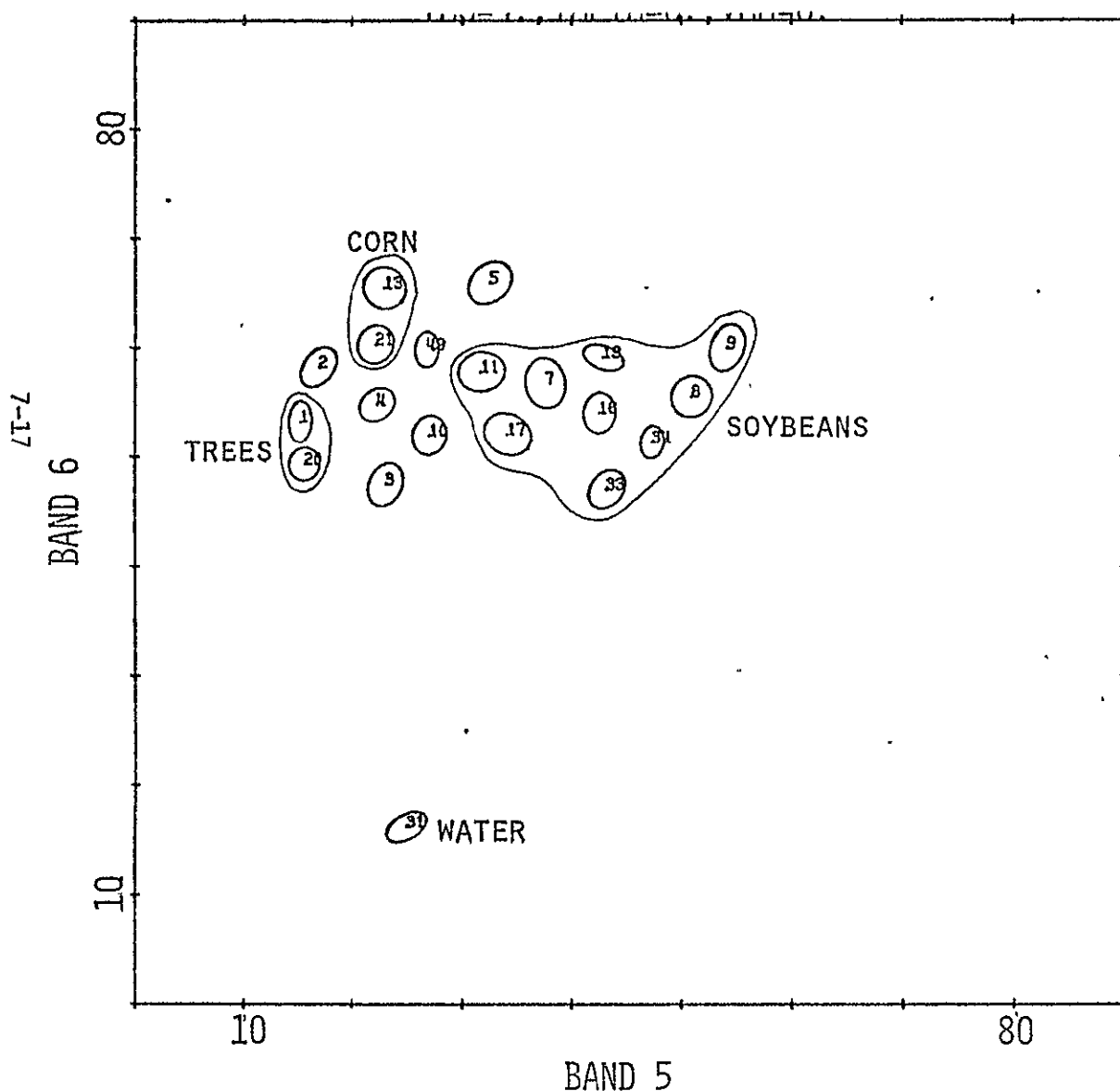
21 AUGUST 1973

CLUSTERS SHOWN CONTAIN 96% OF THE DATA

% OF CROP IN OUTLINED CLUSTERS	CORN	SOY
	80	76

FIGURE 7.8

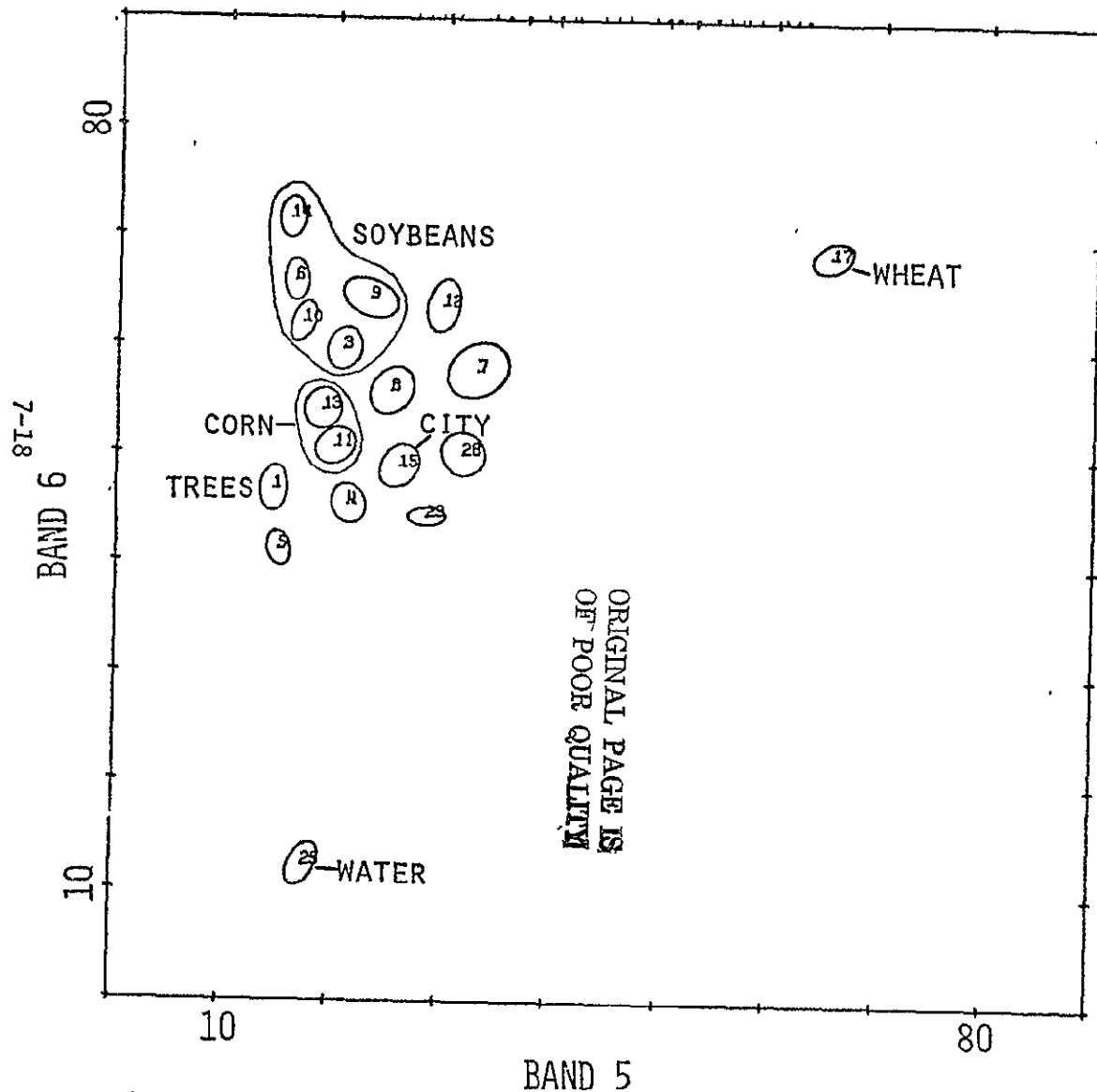
UNITEMPORAL CLUSTER ANALYSIS - FAYETTE COUNTY
RESULTS INCLUDING PIXELS FROM OTHER CROPS, 17 JULY 1973 ACQUISITION



	CLUSTER NAME	
	CORN	SOYBEANS
NUMBER OF PIXELS:	261	513
COMPOSITION (%):		
LABEL CROP	94	78
OTHER CROPS	6	22
% OF MAJOR CROPS:	72	85
OTHER CROPS INCLUDED:	GRAIN (5)	CITY (38)
(NUMBER OF PIXELS)	WHEAT (3)	WHEAT (25)
		BARE (24)
		GRAIN (7)
		WEEDS (3)
1476	TOTAL PIXELS - CLUSTERS SHOWN CONTAIN 85% OF TOTAL	
*NAMING	CRITERION: 60% OF THE CLUSTER CONSISTS OF PIXELS WITH THE GIVEN NAME.	

FIGURE 7.9

UNITEMPORAL CLUSTER ANALYSIS - FAYETTE COUNTY
RESULTS USING ALL LABELED PIXELS, 21 AUGUST 1973 ACQUISITION



	CLUSTER NAME *	
	CORN	SOYBEANS
NUMBER OF PIXELS;	386	474
COMPOSITION (%):		
LABEL CROP	66	87
OTHER CROPS	35	13
% OF MAJOR CROPS:	72	78
OTHER CROPS INCLUDED:	GRAIN (32)	WHEAT (15)
(NUMBER OF PIXELS)	TREES (20)	GRAIN (3)
	WHEAT (11)	
	PASTURE (8)	
	CITY (5)	
	CLOVER (5)	

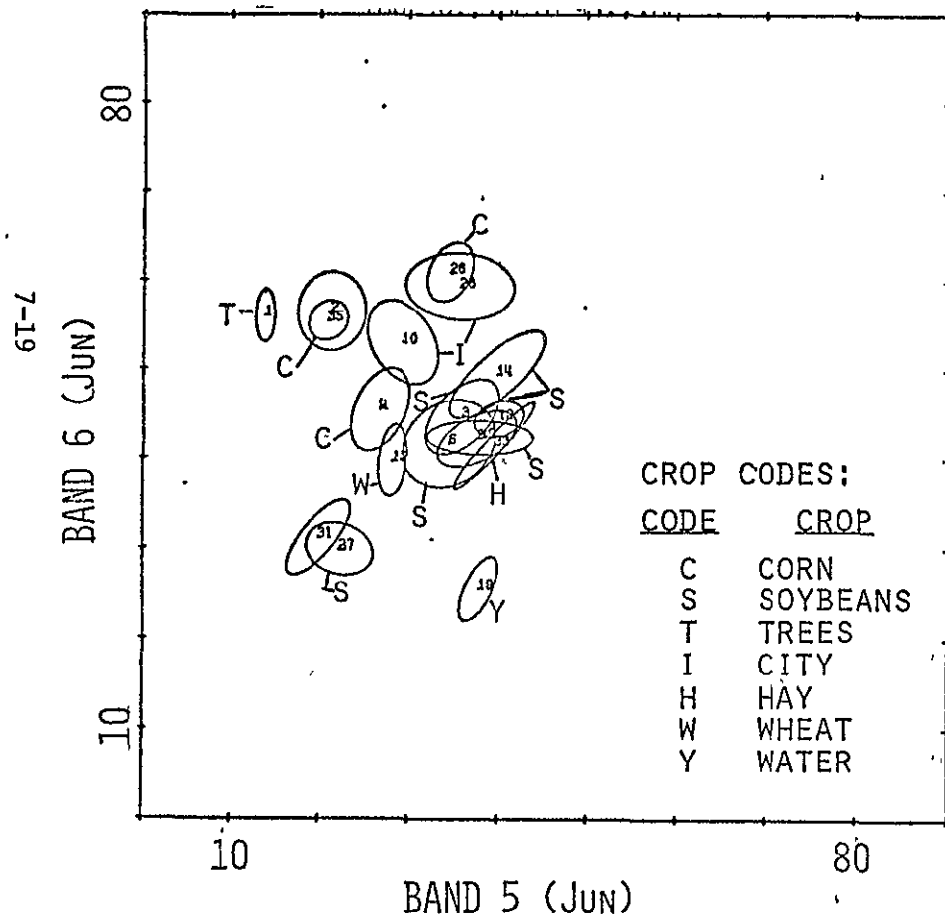
1470 TOTAL PIXELS - CLUSTERS SHOWN CONTAIN
92% OF TOTAL

* NAMING CRITERION: 60% OF THE CLUSTER
CONSISTS OF PIXELS WITH THE GIVEN NAME

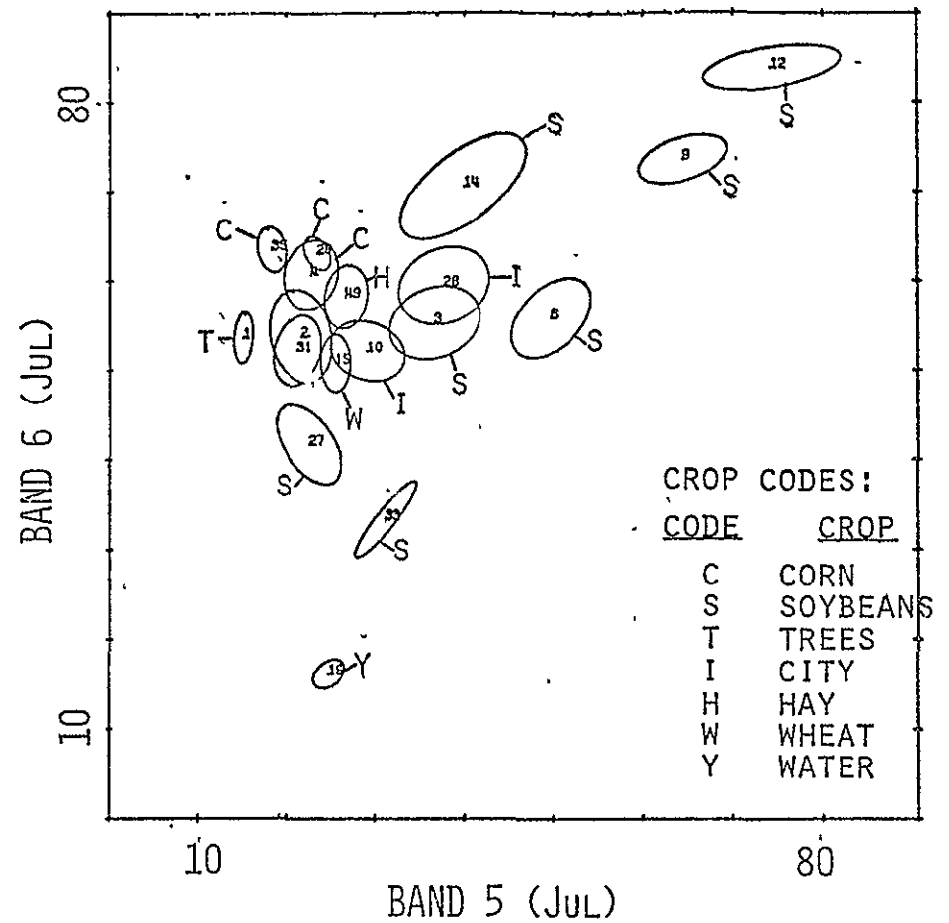
FIGURE 7.10

MULTITEMPORAL CLUSTER ANALYSIS - FAYETTE COUNTY
RESULTS INCLUDING PIXELS FROM OTHER CROPS; THREE DATES, 12 CHANNELS

PATTERNS OF CLUSTERS ON FIRST TWO DATES



29 JUNE 1973



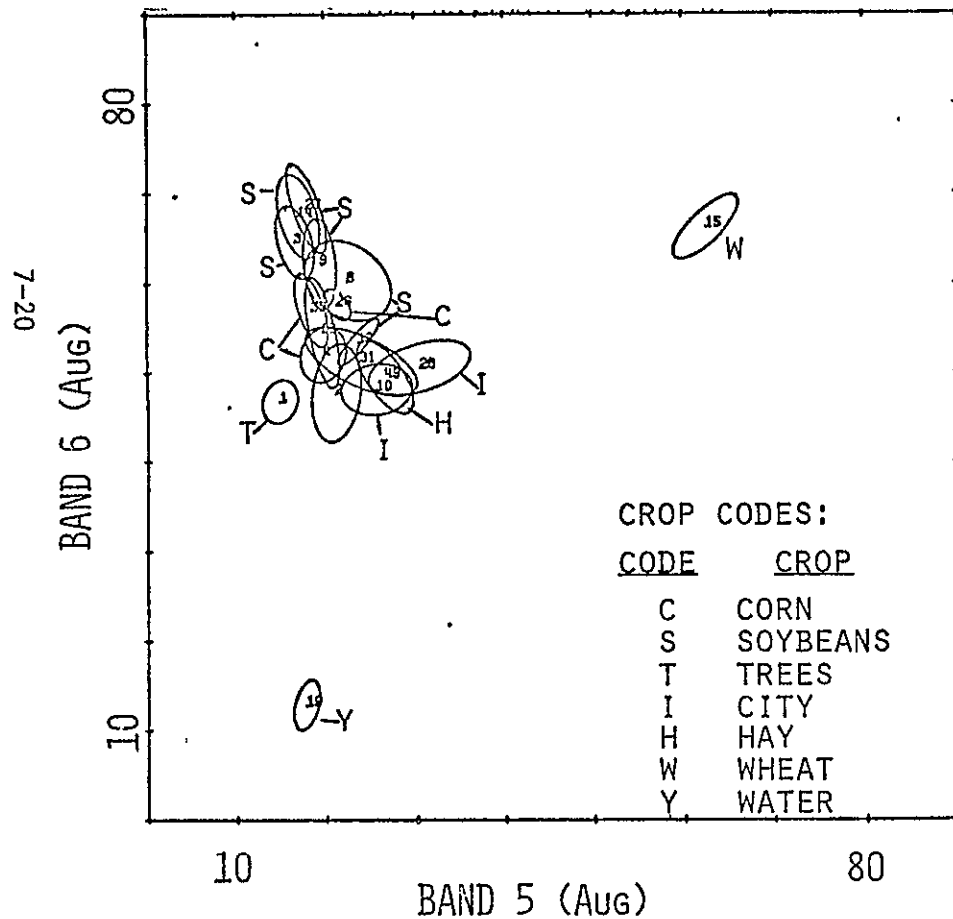
17 JULY 1973

FIGURE 7.11

MULTITEMPORAL CLUSTER ANALYSIS - FAYETTE COUNTY

THREE DATES: 29 JUNE, 17 JULY, 21 AUGUST 1973

RESULTS INCLUDING PIXELS FROM OTHER CROPS (CONTINUED)



CLUSTER STATISTICS

	CLUSTER NAME --	
	CORN	SOYBEANS
NUMBER OF PIXELS:	311	504
COMPOSITION (%):	91	92
LABEL CROP	91	92
OTHER CROPS	9	8
% OF MAJOR CROPS:	84	92

OTHER CROPS INCLUDED: (NUMBER OF PIXELS)	GRAIN (11)	WHEAT (16)
	WHEAT (5)	GRAIN (6)
		BARE (4)

NOTES:

- 1380 PIXELS TOTAL- CLUSTERS SHOWN CONTAIN 92% OF TOTAL
- CROP LABELS OF 21 AUGUST WERE USED

PATTERN OF CLUSTERS ON 21 AUGUST 1973

TASK 8

MULTICROP LABELING AIDS

(R. C. Cicone, P. F. Lambeck, and R. Balon)

8.1 INTRODUCTION

The accuracy of operational large area crop inventory systems modeled after LACIE depends critically on the correctness of crop labels generated by Analyst Interpreters. Task 8 is being conducted as a partial response to a request from the multicrop inventory planning committee for support from the SR&T community in adapting LACIE technology to a multicrop environment. The critical issue addressed is that of analyst labeling.

8.2 OBJECTIVES

Our purpose through Task 8 is to analyze the methods of presentation of Landsat data for purposes of human interpretation to assess how well they convey the information relevant to crop discrimination. We endeavor to develop new data presentation techniques in the form of false color image products as well as graphic displays of spectral information, which stand to expedite correct labeling of crops. Task 8 is divided into two subtasks.

Subtask One has the objective of evaluating and making recommendations for false color image products suited for multicrop labeling purposes.

Subtask Two has the objective of developing and evaluating spectral aids to assist analyst interpretation of multicrop classes.

8.3 APPROACH

Points of approach for this task are layed out in Table 8.1.

8.4 SUMMARY OF PROGRESS FOR THIS QUARTER

The major accomplishments of this quarter are listed in Table 8.2. Four issues were addressed and will be discussed in the following section.

8.5 DESCRIPTION OF PROGRESS

The first issue addressed this quarter was an analysis of the sensitivity of image products to the information content of Landsat data. A method for graphic display of image sensitivity by analytic means was developed during the quarter.

The technical issue under scrutiny is whether the information content in a false color image as viewed by an AI is consistent with the information content of that same scene in digital form. Table 8.3 describes the technique for display of image sensitivity. To begin with, a distance in color space of five counts was empirically determined (using false color image products) to represent a perceptible difference in color of objects. A grid of sample points was then placed in the brightness/greenness plane of the Landsat digital data. This grid was placed at a six count interval and spanned the range of the color production capabilities of the production film converter (i.e., any point outside this range would be saturated in color). One can picture circles centered at each grid point with a radius of three counts and tangent to adjacent circles. These three count circles were chosen so as to conservatively contain the noise characteristics of the Landsat data about the grid points. The question addressed, then, was to determine whether sufficient color resolution was provided, in light of a five count color resolution capability, to preserve information in the Landsat digital data at the level indicated by the circles.

Figure 8.1 illustrates a typical result using Product 1 color mapping procedures. Each ellipse represents visual sensitivity to color changes about the center. Resolution of the data into color is poor along the direction of elongation. Smaller ellipses mean the data space is better represented by separable colors. Large ellipses indicate that color varies slowly. It is certainly clear that three count tangent circles are not preserved in this mapping. Figure 8.1 indicates poor resolution of the data into color and distortion of the information

in certain portions of the space. Figure 8.2 overlays Figure 8.1 to show where principal crop distributions corresponding to this acquisition fall.

We have found that the characteristic pattern of sensitivity results from the exposure control linkage of the Production Film Converter, which relates input level exponentially to film transmission. This results in decreasing sensitivity toward less than the highest input levels, or, equivalently, the highest Landsat channel values. Overall image sensitivity is very much a function of the bias and scale factors selected for data transformation before image generation. It was discovered that the image sensitivity resulting from use of transformation factors appropriate to an acquisition displaying the full natural range of agricultural data, was inadequate for resolving differences of less than 7 data counts.

A hypothetical PFC calibrated to relate input level to film transmission linearly (rather than exponentially) was investigated. Image sensitivity was found to be lower. Figure 8.3 shows the sensitivity display produced in this experiment.

A second study conducted this quarter was a probability of misclassification analysis of crop distributions processed through various image generation procedures. This was done with the idea of determining if and to what extent loss of information occurred in the data-to-false-color-image transformation. The crop distributions used were wheat and non-wheat clusters for the Kansas sample segment 1154, acquisition of day 124, 1976 (Figure 8.2). This acquisition was chosen for the good separation between these two classes.

We found that dropping a channel of spectral information (Landsat Channel 3) caused a sizeable increase in expected misclassification between the crop distributions. Table 8.4 presents details. This has been recognized empirically and the AI is routinely provided a product which includes Channel 3 (Product 2). However, the price is that this

image does not contain Channel 1 information. We expect that much of the information of Landsat data lies in the interrelationship of the four channels considered simultaneously. For this reason, interpreting two images with subsets of bands is not the same as interpreting one image with all the information. Data feature extraction transformations exist which reduce the dimensionality of Landsat data without significant loss of information. The linear Kauth transformation is a prime example. There is impetus for investigating ways of converting Landsat features to a single color image with full retention of information.

The two crop distributions were converted into corresponding color distributions, simulating the Product One algorithm and proposed variations as well. Tables 8.5 and 8.6 present details. In each case an increase of confusion between the distributions was observed which was of comparable seriousness to that due to dropping a band. This confusion is attributable to incomplete resolution of the information in the data into color. We have examined PFC products for which color coordinates of field means have been computed. From this we judge the human eye cannot well distinguish colors separated by 5 color space units or less (we use the 1976 CIE Uniform Color Space). For the present we use 5 units as our standard of color resolution. If an image product takes meaningful data variation into perceptually indistinguishable colors then information is lost. Our sensitivity analysis presented above suggested this was happening for a typical scene, and our classification accuracy study tends to confirm that conclusion.

Our third activity during the quarter related to facilitating the transition to multicrop interpretation and labeling. The variation of corn and soybeans over time was plotted in color space coordinates. The data used was from CITARS acquisitions of Fayette County. Figure 8.4 shows color trajectories of the crops in the principal components of color variation.

A fourth activity, in response to reports of apparent problems in applying the SCREEN and XSTAR algorithms to CITARS data, determined that the apparent problems observed by ERIM and by UCB were caused in each case by errors in processing. Our analysis indicates that the Landsat-1 to Landsat-2 data conversion previously specified by ERIM is actually very close to the optimum conversion for CITARS data. SCREEN and XSTAR performed normally on this data.

8.6 CONCLUSIONS AND RECOMMENDATIONS

Specific conclusions are listed in Table 8.7. We recommend investigation of new data-to-image processing to minimize loss of information content.

8.7 PLANS

Plans for the coming quarter are detailed in Table 8.8.

Table 8.1

APPROACH

SUBTASK 1

- USE PRINCIPLES OF COLOR SCIENCE TO MODEL THE COLOR PRODUCTION CHARACTERISTICS OF THE PFC
- ANALYZE TECHNIQUES FOR PRODUCING FALSE COLOR IMAGERY FROM LANDSAT DATA IN TERMS OF:
 - SPECTRAL INFORMATION CONTENT
 - VARIATION OF COLOR SENSITIVITY IN IMAGE DATA SPACE
- EVALUATE TRENDS OF COLOR VARIATION FOR CORN AND SOYBEANS AS FUNCTION OF DATE
- EXAMINE ALTERNATIVE SINGLE AND MULTI-DATE IMAGE PRODUCTS

SUBTASK 2

- EVALUATE EXISTING MULTICROP DATA SETS TO EXAMINE:
 - EFFECTS OF HAZE AND SUN ANGLE COMPENSATION TECHNIQUES
 - DEVELOP SPECTRAL AIDS PARALLEL TO THOSE CURRENTLY EMPLOYED IN LACIE
- EXAMINE UTILITY OF HAZE AND SUN ANGLE-CORRECTED IMAGE PRODUCTS

Table 8.2

SUMMARY OF RESULTS

- A METHOD FOR GRAPHIC DISPLAY OF SENSITIVITY OF FILM PRODUCTS THROUGHOUT DATA SPACE WAS DEVELOPED AND USED TO EVALUATE IMAGE SENSITIVITY.
- AN ANALYSIS OF IMAGE PRODUCTS WAS PERFORMED IN TERMS OF PROBABILITY OF MISCLASSIFICATION OF CROP DISTRIBUTIONS IN A COLOR SPACE THAT IS PERCEPTUALLY UNIFORM WITH RESPECT TO THE HUMAN EYE.
- PREPARATION OF CITARS DATA SET.
- GROWING SEASON COLOR TRENDS FOR CORN AND SOYBEAN WERE EXAMINED USING CITARS DATA.
- PERFORMANCE OF THE SCREEN AND XSTAR ALGORITHMS ON CITARS DATA WAS VERIFIED.

Table 8.3
IMAGE SENSITIVITY DISPLAY PROCEDURE

- SAMPLE POINTS IN DATA PLANE AT A SPACING OF 6 COUNTS
- TRANSFORM THEM TO UNIFORM COLOR SPACE
- CONSTRUCT A SYMMETRIC NORMAL DISTRIBUTION ABOUT EACH POINT WITH A STANDARD DEVIATION OF 5 UNITS
- ANALYTICALLY TRANSFORM EACH DISTRIBUTION BACK TO THE DATA PLANE AND DISPLAY

Table 8.4

INFORMATION LOSS DUE TO USING A SUBSET OF FEATURES
DURING BIOWINDOW 2 (S/S 1154)

<u>VARIABLES</u>	<u>PERCENT MISCLASSIFICATION</u>	
	<u>WHEAT</u>	<u>OTHER</u>
BANDS 1, 2, 3 AND 4	2.8	5.4
BANDS 1, 2, AND 4	12.2	4.8
TASSELLED CAP BRIGHTNESS/GREEN	10.4	4.6

Table 8.5
 INFORMATION LOSS DUE TO COLOR RESOLUTION OF THE HUMAN EYE

		<u>PERCENT MISCLASSIFICATION</u>	
		<u>WHEAT</u>	<u>OTHER</u>
• PRODUCT 1	$\sigma = 0$	15.2	5.0
• PRODUCT 1	$\sigma = 5$	18.2	8.4

Table 8.6
INFORMATION LOSS DUE TO FILM PRODUCT
BIOWINDOW 2 (S/S 1154)

<u>PRODUCT AND PROCESSING</u>	<u>PERCENT MISCLASSIFICATION</u>	
	<u>WHEAT</u>	<u>OTHER</u>
• NO PREPROCESSING		
PRODUCT 1	18.2	8.4
KANEKO PRODUCT	18.4	17.4
KRAUS PRODUCT	26.4	32.8
• DOUBLE BAND 4		
KANEKO PRODUCT	17.2	6.2
HOCUTT PRODUCT	17.2	6.6
• KANEKO'S PROCESSING		
KANEKO PRODUCT	20.6	8.6
HOCUTT PRODUCT	18.6	6.4
KRAUS PRODUCT	21.2	9.6

Table 8.7
CONCLUSIONS

- WITH BAND 4 DOUBLING, KANEKO, HOCUTT AND PRODUCT 1'S RETAIN COMPARABLE INFORMATION AND DISPLAY SIMILAR COLOR SENSITIVITY
- KANEKO PREPROCESSING RESULTED IN SLIGHT DETERIORATION OF EXPECTED CLASSIFICATION USING ALTERNATIVE PRODUCTS; KRAUS PRODUCT LAGS IN THE COMPARISON
- NO PREPROCESSING RESULTED IN THE POOREST EXPECTED CLASSIFICATION
- WITH EITHER EXPONENTIAL OR LINEAR FILM TRANSMISSION CONTROL, COLOR RESOLUTION OF THE EYE MASKS AVAILABLE RESOLUTION IN THE DATA IN THE REGION OF DATA CONCENTRATION
- BIAS AND SCALING TECHNIQUES UTILIZING BANDS 1, 2, AND 4 AS INPUTS TO THE PFC COLOR GUNS REDUCES INFORMATION CONTENT OF THE DATA AND DOES NOT PRESERVE DISTANCE RELATIONSHIPS IN THE DATA

Table 8.8

PLANS

- PREPARE AN INTERIM TECHNICAL REPORT
- EXAMINE IMAGE-TO-IMAGE COLOR CHARACTERISTICS OF CROPS
- DEVELOP AND EVALUATE TECHNIQUES FOR DATA-TO-IMAGE MAPPING WHICH WILL RETAIN INFORMATION CONTENT OF IMAGERY
- CONSIDER MULTITEMPORAL IMAGE PRODUCTS
- BEGIN DEVELOPMENT OF AI SPECTRAL AIDS BASED ON CURRENT RESULTS FROM CROP DISCRIMINABILITY TASK
- INVESTIGATE APPLICATION OF XSTAR HAZE CORRECTION TO ADDRESS LOCALIZED ATMOSPHERIC VARIATIONS WITHIN THE BOUNDARIES OF LACIE SEGMENTS

Figure 8.1

SEGMENT 1154 BIO 2
PRODUCT ONE

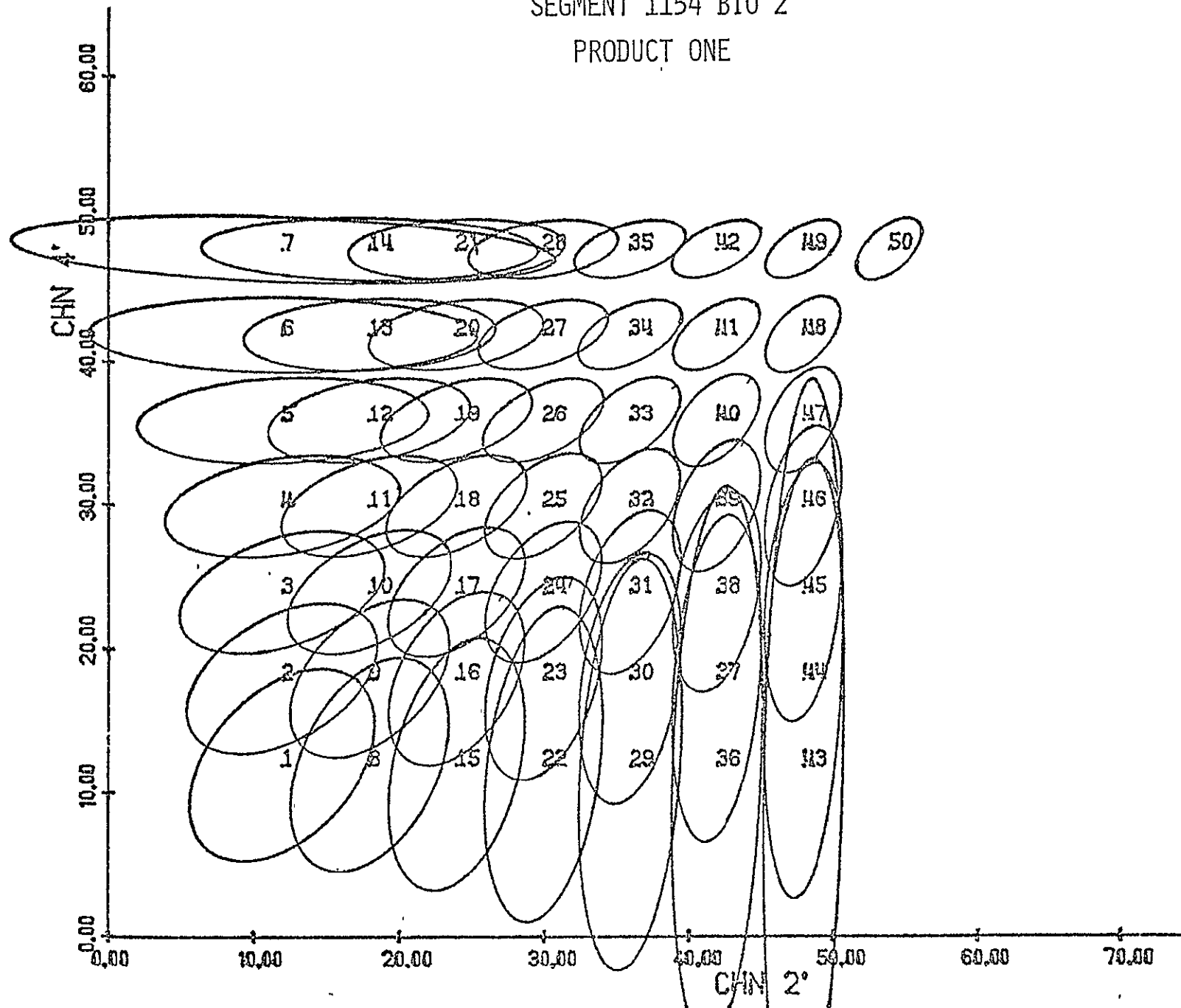
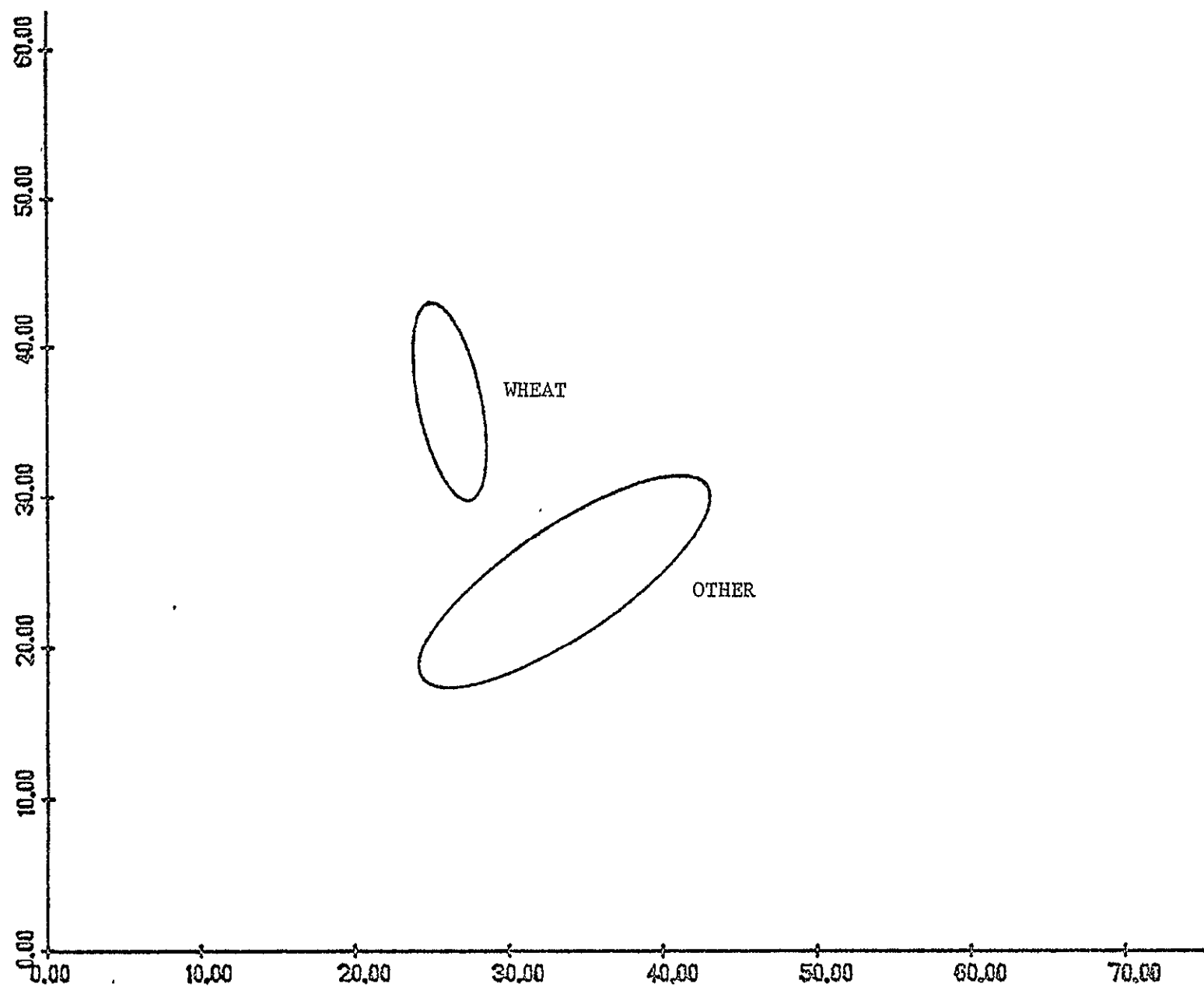


Figure 8,2

Data Location for S/S 1154 acquisition
in Bio 2



8-15

Figure 8.3
 SEGMENT 1154 BIO 2
 PRODUCT ONE WITH LINEAR PFC CONTROL

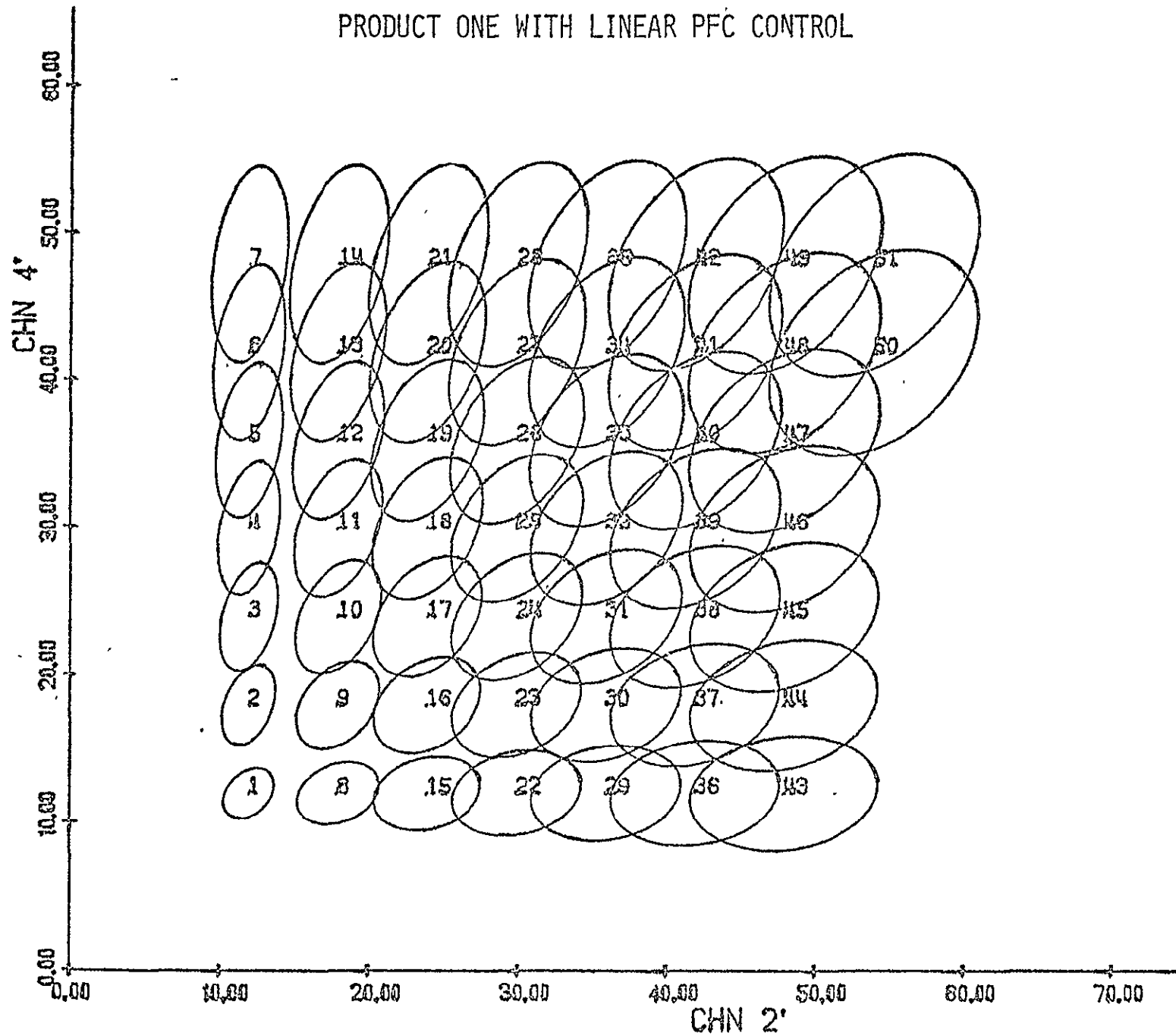


Figure 8.4 (a)
CORN COLOR TRAJECTORY
UNIFORM COLOR SPACE

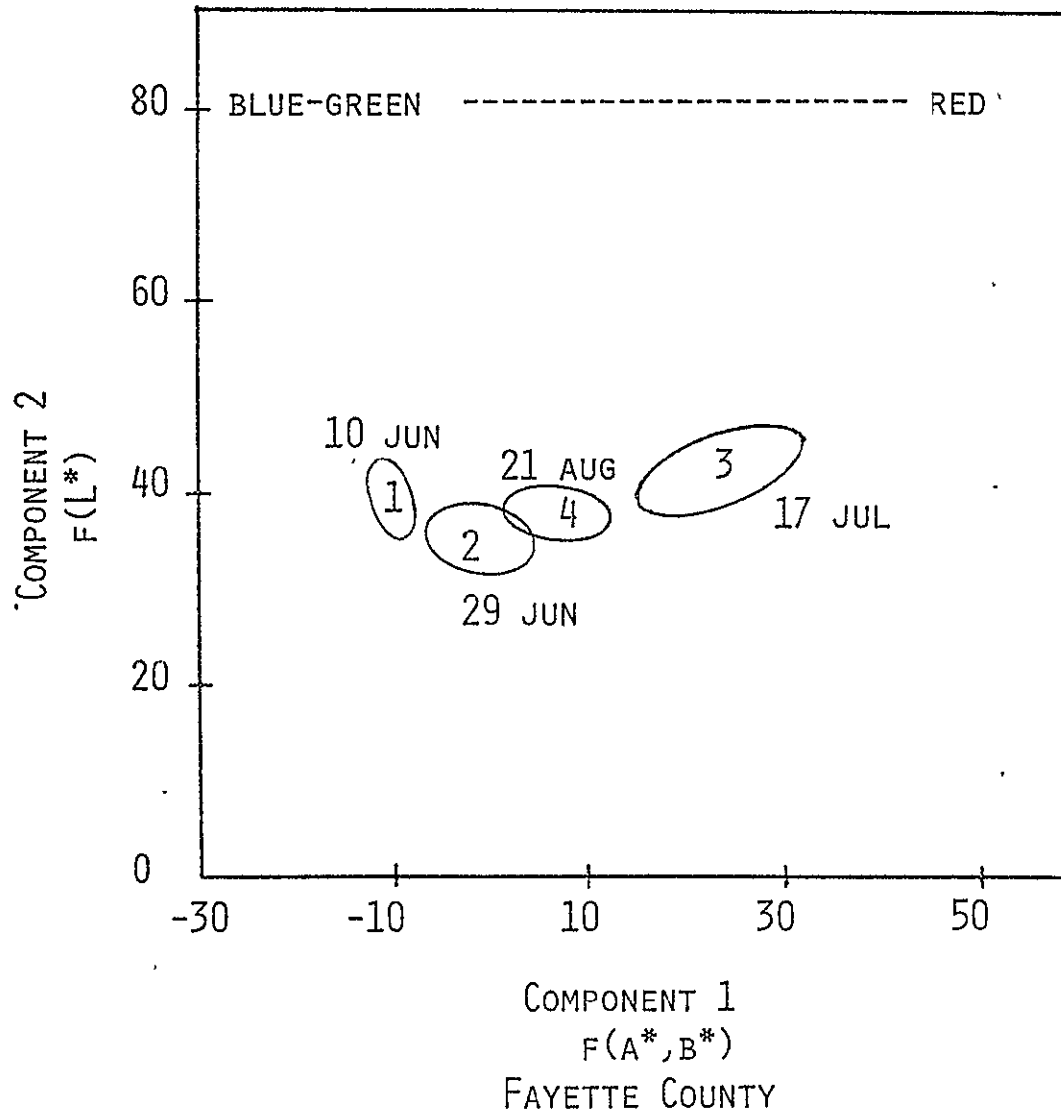


Figure 8.4 (b)
 SOYBEAN COLOR TRAJECTORY
 UNIFORM COLOR SPACE

